Pivoting algorithms for complementarity problems in economics

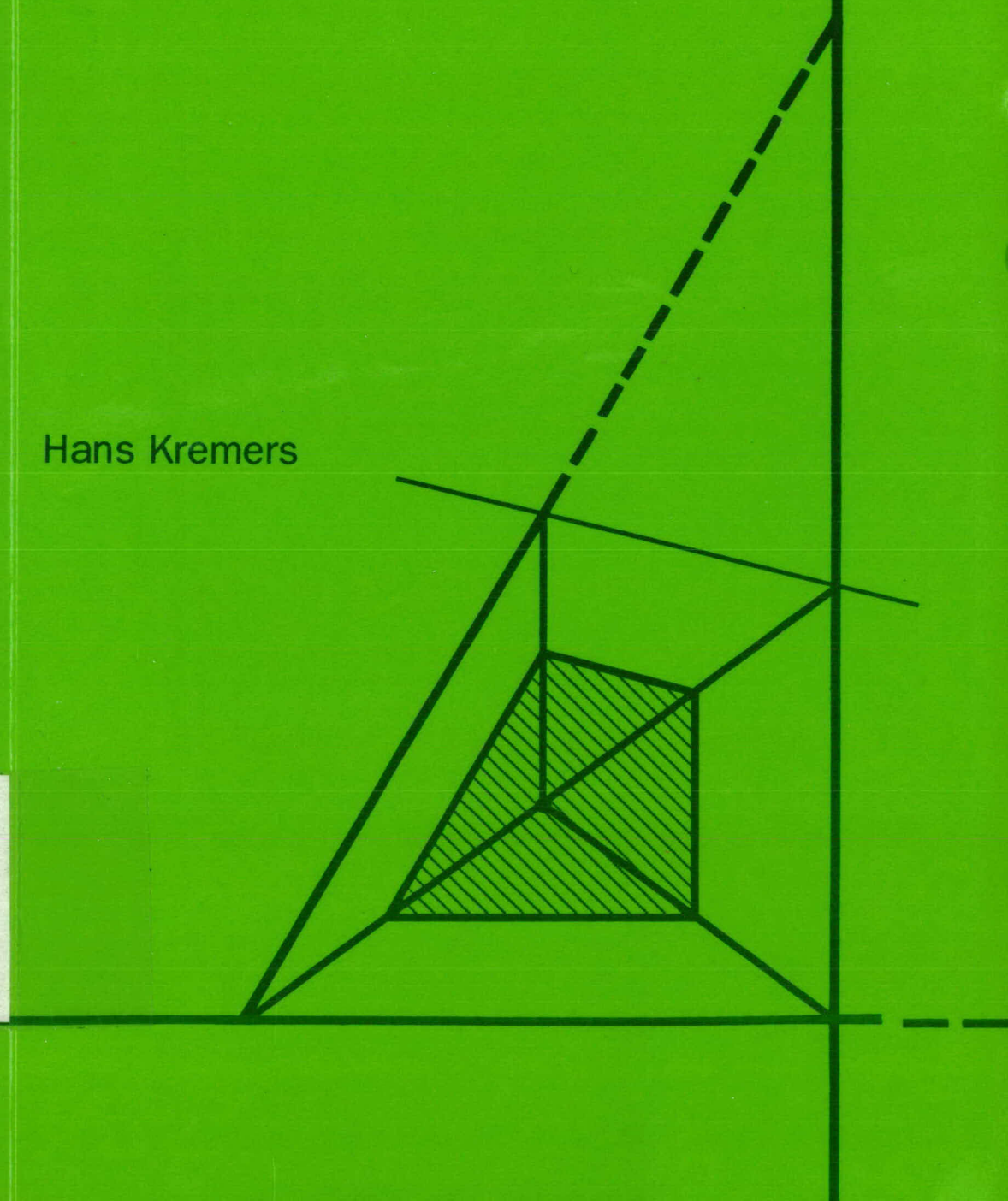
Kremers, Johannes Antonius Wilhelmus Maria

Publication date:
1991

Document Version
Publisher's PDF, also known as Version of record

Pivoting Algorithms for Complementarity Problems in Economics

Hans Kremers



Pivoting Algorithms for Complementarity Problems in Economics

Proefschrift

ter verkrijging van de graad van doctor aan de Katholieke Universiteit Brabant, op gezag van de rector magnificus, prof. dr. L.F.W. de Klerk, in het openbaar te verdedigen ten overstaan van een door het college van dekanen aangewezen commissie in de aula van de Universiteit op vrijdag 20 december 1991 te 16.15 uur

door

Johannes Antonius Wilhelmus Maria Kremers

geboren te Roosendaal c.a.

druk: wibro dissertatiedrukkerij, helmond.

PROMOTOR: Prof. dr. A.J.J. Talman

Acknowledgement

This monograph contains the results of four years of research carried out at the Department of Econometrics of Tilburg University. This research has been done in cooperation with Dolf Talman. I am very much indebted to him for introducing me to the field of simplicial algorithms and complementarity problems in economics but most of all I thank him for being more than just a supervisor. Furthermore I would like to express my appreciation to Pieter Ruys for creating an encouraging atmosphere for doing research which is unique for mathematical economists in the Netherlands.

I wish to thank Professor Joaquim Júdice of Coimbra University, Portugal, Professor Gerard van der Laan of Free University Amsterdam, Dr. Hans van Maaren of Delft University, and Professors Stef Tijs, Pieter Ruys, and Dolf Talman of Tilburg University for taking part in my Ph.D.-defense committee and for their remarks on the thesis. I also would like to thank Willy Spanjers for being my roommate for the last three years and for the many encouraging discussions concerning economics and much more I had with him during these three years.

I thank Jean-Jacques Herings, Gerard van der Laan, and Dolf Talman for the meticulous proofreading of the manuscript. The many drawings in this manuscript have been made by Yvonne van Delft and Jan Pijnenburg. Finally, I would like to thank all my friends and colleagues who made my stay at Tilburg University a very pleasant one.

Hans Kremers
November 1991

Aan mijn ouders

Contents

Part A

Preliminaries

Chapter 1

Introduction

The purpose of this monograph is twofold. On the one hand algorithms are introduced to compute an equilibrium in an economy with linear production technologies. On the other hand several new algorithms are presented to compute a solution to complementarity problems, well-known from mathematical programming. These applications have a very close relationship as most economic equilibrium problems take the form of some complementarity problem.

All the algorithms presented in this monograph are path-following algorithms, i.e. they start in some arbitrarily chosen point in some given convex set and follow a path of points in that set either towards a solution of the problem or towards infinity. The pioneering work on solving equilibrium problems by applying path-following algorithms is due to Scarf. In Scarf (1967) and Scarf (1973) so-called fixed point algorithms were proposed to compute economic equilibria. A fixed point of a function is a point being its own image under this function. In 1910 the Dutch mathematician L.E.J. Brouwer proved that any continuous function from a compact, convex, nonempty set into itself has at least one fixed point (see Brouwer (1912)). Scarf utilized an argument of Lemke and Howson (1964) and Lemke (1965) and the relationship of Sperner's lemma (see Sperner (1928)) to the Brouwer fixed point theorem to introduce a path-following algorithm on the unit simplex. In a subdivision of the unit simplex into so-called primitive sets the algorithm finds within a finite number of iterations a primitive set containing an approximation of a fixed point. Starting in a corner of the unit simplex Scarf's algorithm generates a sequence of

adjacent primitive sets in this subdivision. In Hansen (1968) this algorithm was improved by using a subdivision of the unit simplex into simplices. This simplicial subdivision was discovered by Freudenthal (1942) and made operational on the unit simplex by Kuhn (1960). The advantage of a simplicial subdivision is that it can be implemented very easily to follow a sequence of adjacent simplices in a simplicial subdivision.

In van der Laan and Talman (1979) and van der Laan and Talman (1981) simplicial variable dimension restart algorithms were introduced to compute fixed points. These algorithms were capable of starting in an arbitrarily chosen grid point of the underlying simplicial subdivision and they generate a sequence of adjacent simplices of varying dimension through this simplicial subdivision towards a simplex containing an approximating solution. Simplicial variable dimension restart algorithms were quite an improvement on preceding simplicial algorithms which were either stuck to a fixed starting point such as in Kuhn (1960) or needed an embedding in a set having an extra dimension such as the homotopy method in Merrill (1971) or Eaves (1972).

In Doup and Talman (1987) the algorithms introduced by van der Laan and Talman were improved by introducing a new type of simplicial subdivision. This new simplicial subdivision is known as the V-triangulation. An advantage of the V-triangulation is that this simplicial subdivision depends on the chosen starting point so that one is not restricted to grid points of a predetermined simplicial subdivision anymore. The introduction of the V-triangulation by Doup and Talman led to the ability to solve equilibrium problems in pure exchange economies in such a way that the sequence or path followed by the algorithms can be interpreted as a tâtonnement process in which prices are simultaneously adjusted (compare Walras (1874)). In Doup (1988) several algorithms were presented being capable of computing an approximating equilibrium of an exchange economy for any finite accuracy.

The ideas behind the V-triangulation paved the way to succesfully attacking problems like complementarity problems and equilibrium problems in pure exchange economies. The flexibility of the V-triangulation made it possible to develop efficient algorithms for each problem under consideration instead of rewriting the problem to one that can be solved by an existing algorithm.

An economy with linear production technologies consists of commodities, consumers endowed with each of these commodities and demanding a certain amount of these commodities depending on their prices, and activities whose production levels determine the amount of output or input of each commodity in the economy. An equilibrium in such an economy is given by a vector of prices of the commodities and activity levels such that for each commodity total demand is met by total supply and no activity is making positive profit. One of the main contributions to the computation of equilibria in an economy with linear production thusfar is the algorithm proposed in Mathiesen (1985b). Mathiesen proposed to solve the system of equilibrium conditions in an economy with linear production technologies by a sequence of linear complementarity problems (*SLCP*).

First Mathiesen observed that prices should be normalized on the unit simplex. Then he chose an arbitrary price vector in the unit simplex to serve as a starting point to the *SLCP* algorithm. In this arbitrarily chosen price vector Mathiesen calculated the so-called first-order Taylor-expansion of the economy's excess demand function thereby changing the system of nonlinear equilibrium conditions into a linearized system of equations.

The problem that because of homogeneity of degree zero of demand the Jacobian matrix of the demand function, being the matrix with its first-order derivatives, is singular is traditionally solved in either of two ways. One can add a normalization constraint on prices, e.g., taking prices only on the unit simplex on which they sum up to one, or one can stipulate a numéraire price being fixed on some arbitrarily positive value. Mathiesen (1985b) chose the latter approach thereby allowing the algorithm to generate prices outside the unit simplex when solving the linearized system.

The linearized system takes the form of a linear complementarity problem. By applying the Lemke algorithm (1965) Mathiesen either could find a solution to the linear complementarity problem or he could not even when a solution to the linear complementarity problem does exist. If no solution could be found given the numéraire choice, Mathiesen suggested just to take another price as a numéraire and to repeat the Lemke procedure for the linear complementarity problem then obtained, until some linear complementarity problem has been found where applying the Lemke algorithm results in a solution. The price vector resulting from the Lemke algorithm

is then projected on the unit simplex and taken as a starting point to the next iteration. In this way a sequence of price vectors on the unit simplex is generated which possibly converges to an equilibrium solution.

Eaves (1987) suggested to escape the impasse caused by the singular Jacobian matrix of the demand function by applying the first approach mentioned above, i.e., adding a normalization constraint on the prices. Furthermore Eaves proved that the equilibrium problem in an economy with linear production technologies is equivalent to the stationary point problem of the excess demand function on a subset of the unit simplex defined by the "no-profit"-conditions on the equilibrium prices. Then Eaves rewrote the problem in such a way that the stationary point problem can be solved by a sequence of linear complementarity problems in a way similar to Mathiesen (1985b). According to Eaves this alternative appears to lend itself more easily to theoretical analysis than Mathiesen's *SLCP*. Furthermore the linear complementarity problem in each iteration of Eaves' *SLCP* is proved to have a solution in each iteration contrary to Mathiesen (1985b).

Kamiya and Talman (1990) introduced a very efficient algorithm to find a stationary point of an affine function on a polytope. Recalling that the equilibrium problem in an economy with linear production technologies is equivalent to the stationary point problem of the excess demand function on a subset of the unit simplex being a polytope, we suggest to approximate the latter problem by a sequence of linear stationary point problems in that set and solve each linear stationary point problem in this sequence by an algorithm based on the ideas of Kamiya and Talman (1990). In this way an alternative to Eaves (1987) and Mathiesen (1985b) is obtained which appears to be much more efficient.

The algorithms to solve the equilibrium problem in an economy with linear production technologies by approximating the problem by a sequential algorithm seem to be quite fast if they are able to generate a converging sequence of approximating solutions. The main drawback of this kind of algorithms is however that global convergence of the obtained sequence cannot be proved. To overcome this problem we also introduce an alternative to this kind of algorithms. We will introduce a simplicial variable dimension restart algorithm in order to solve the equilibrium problem. It approximates the path of prices resulting from the adjustment process introduced in

van den Elzen, van der Laan, and Talman (1990) by linearizing the excess demand function on a simplicial subdivision of the unit simplex. As this adjustment process is proved to converge to an equilibrium solution the simplicial algorithm introduced in this monograph ends up with an approximating solution. If one is not satisfied with the approximation the procedure can be restarted with a finer simplicial subdivision of the unit simplex. Contrary to the sequential algorithms this kind of algorithm might be rather slow once a high accuracy of the approximation has been obtained.

The equilibrium problem in an economy with linear production technologies takes the shape of a complementarity problem. This is a problem well-known from mathematical programming as well as from other areas of research where optimization plays an important role. Complementarity problems are the subject of another part of this monograph. The simplest and probably most renowned complementarity problem among them is the linear complementarity problem. This problem is to find nonnegative vectors which are affinely related and complementary. The linear complementarity problem contains well-known problems such as linear programming, quadratic programming, and the equilibrium problem in bimatrix games as special cases. Furthermore the problem is frequently met when solving Karush-Kuhn-Tucker conditions. The most well-known algorithm to compute a solution to the linear complementarity problem is the Lemke algorithm introduced in Lemke (1965) and Lemke and Howson (1964). The Lemke algorithm is also known as the complementary pivoting algorithm. One of its main drawbacks is however that one is stuck to a fixed starting point, causing difficulties in performing sensitivity analysis or inefficiencies in restarting when solving a sequence of linear complementarity problems. In Talman and Van der Heyden (1983) a class of algorithms was introduced generalizing the Lemke complementary pivoting algorithm to the incorporation of an arbitrarily chosen starting point. In this monograph an alternative algorithm incorporating the possibility of an arbitrarily chosen starting point is introduced being more efficient and having a more natural interpretation than the algorithms introduced by Talman and Van der Heyden.

In Cottle (1966) the nonlinear complementarity problem was introduced where the affine relation between the vectors is replaced by a continuous one. The equilibrium problems in pure exchange economies and exchange economies with linear production

technologies all take the form of a nonlinear complementarity problem. Hence the algorithms introduced by Mathiesen (1985b) can also be used to find a solution to the nonlinear complementarity problem. Kojima (1974), Fisher and Gould (1974), Garcia (1973), and Merrill (1971) for example introduced algorithms to solve the nonlinear complementarity problem. All these methods originated as a variant of the algorithm introduced in Scarf (1967) in constructively proving the Brouwer fixed point theorem. In this monograph we introduce a simplicial variable dimension restart algorithm to compute a solution to the nonlinear complementarity problem.

The nonlinear complementarity problem is a generalization of the linear complementarity problem to an arbitrary continuous function as the relation between the two vectors in the problem. The next step is to put arbitrary upper and lower bounds on the variables. This problem is referred to as the generalized nonlinear complementarity problem or the nonlinear complementarity problem with lower and upper bounds. In van der Laan and Talman (1985) an algorithm is introduced to solve the linear case while in van der Laan and Talman (1987) the nonlinear case is solved. Van der Laan and Talman (1987) use the K'-triangulation (see Todd (1978)) to underly the algorithm. This triangulation is based on the one introduced by Freudenthal (1942). In this monograph a simplicial variable dimension restart is introduced which makes use of the V-triangulation.

This monograph consists of three parts. Part A is an introduction to the rest of the monograph. In Chapter 2 the mathematical preliminaries necessary for understanding the algorithms in this monograph are introduced as well as the notation used throughout the monograph. It contains some basic theorems and results from linear algebra concerning polytopes and cones and it shows the relation between fixed points, stationary points, and complementarity. Chapter 3 considers the basics underlying the algorithms introduced in this monograph and introduces a general framework underlying all the algorithms introduced in this monograph. Among other things it describes the pivoting procedure well-known from linear programming. Furthermore it introduces the concept of a simplicial subdivision of a convex set.

In Part B algorithms to solve well-known complementarity problems are introduced. Chapter 4 describes the Lemke complementary pivoting algorithm as well

as a slightly adapted version of one of the algorithms introduced by Talman and Van der Heyden using the framework of Chapter 3. Furthermore it discusses some familiar problems in mathematical programming being a special kind of the linear complementarity problem. Chapter 5 introduces a new algorithm to solve the linear complementarity problem incorporating the possibility of an arbitrary starting point. This chapter is based on Kremers and Talman (1990c). Chapter 6 of this monograph considers the nonlinear complementarity problem with lower and upper bounds and introduces a variable dimension restart algorithm to find a solution. It extends the ideas obtained from the V-triangulation to a new simplicial subdivision of the cube to underly the algorithm. This chapter is based on Kremers and Talman (1990a). Chapter 7 extends the algorithm introduced in Chapter 6 to the nonlinear complementarity problem. This chapter is based on Kremers and Talman (1990b).

Part C of this monograph concentrates on algorithms to compute an equilibrium in an economy with linear production technologies. The concept of an economy with linear production technologies as well as the *SLCP*-algorithms of Mathiesen (1985b) and Eaves (1987) are introduced in Chapter 8. Chapter 9 presents the sequence of linear stationary point problems to approximate the equilibrium problem in an economy with linear production technologies as an alternative to the algorithms introduced in Chapter 8. This chapter is based on Kremers and Talman (1991). Chapter 10 introduces a simplicial variable dimension restart algorithm to follow the adjustment process introduced in van den Elzen, van der Laan, and Talman (1990). This chapter is based on van den Elzen, Kremers, van der Laan, and Talman (1991) and closes this monograph.

Chapter 2

Mathematical preliminaries

This chapter establishes the necessary mathematical background for the understanding of this monograph. We assume however some familiarity of the reader with basic linear algebra. In Section 1 we introduce some basic notation and basic concepts used throughout the monograph. In Section 2 of this chapter we discuss properties of a specific kind of convex sets being frequently met throughout this monograph. For a more detailed understanding of the material introduced in that section we refer to Schrijver (1986).

All the problems presented in this monograph have the structure of a stationary point problem. Section 3 of this chapter defines a stationary point problem and describes its relation with fixed points and complementarity. Also the existence of solutions to complementarity problems and stationary point problems is proved in that section, mainly by an extensive use of the results obtained from fixed point theory.

2.1 Basic concepts and notations

Throughout this monograph $\mathbf{R}$ denotes the set of real numbers and $\mathbf{N}$ the set of positive integers. For $n \in \mathbf{N}$, the subset $\mathcal{I}_n$ of $\mathbf{N}$ denotes the set $\{1, \ldots, n\}$ and $\mathbf{R}^n$ the n-dimensional Euclidean space. An element or point x in $\mathbf{R}^n$ is also called an n-dimensional vector or n-vector and consists of n real numbers, denoted x_i, $i \in \mathcal{I}_n$, as its components in such a way that they make up a column vector. The n-vector $x^\top$, denoting the transpose of x, represents x in $\mathbf{R}^n$ as a row vector, i.e. $x^\top = (x_1, \ldots, x_n)$.

For $i \in \mathcal{I}_n$, the n-vector $e(i)$ in $\mathbf{R}^n$ with i-th component equal to one and all other components equal to zero is called the i-th unit vector in $\mathbf{R}^n$. The n-vector e in $\mathbf{R}^n$ denotes the n-vector with all components equal to one.

An n-vector s in $\mathbf{R}^n$ with only components $s_i \in \{-1, 0, +1\}$ for $i \in \mathcal{I}_n$ is called a *sign vector*. Furthermore, for $x \in \mathbf{R}^n$ the n-vector $\mathrm{sgn}(x)$ is a sign vector such that $\mathrm{sgn}_i(x) = +1$ if $x_i > 0$, $\mathrm{sgn}_i(x) = 0$ if $x_i = 0$, and $\mathrm{sgn}_i(x) = -1$ if $x_i < 0$, for all $i \in \mathcal{I}_n$. Given such a sign vector $s \in \mathbf{R}^n$ the sets $\mathcal{I}^{+1}(s)$, $\mathcal{I}^0(s)$, and $\mathcal{I}^{-1}(s)$ are defined as $\{j \mid s_j = +1\}$, $\{j \mid s_j = 0\}$, and $\{j \mid s_j = -1\}$ respectively.

The n-vectors $x^1, \ldots, x^r$ in $\mathbf{R}^n$ can be put into an $(n \times r)$-matrix X such that these vectors form the columns of the matrix X. In this monograph $X_{\cdot j}$ or x^j denotes the j-th column of an $(n \times r)$-matrix X while $X_{i\cdot}$ denotes the i-th row of this matrix, i.e. $X_{\cdot j} = (x_1^j, \ldots, x_n^j)^\top$ for $j \in \mathcal{I}_r$ and $X_{i\cdot} = (x_i^1, \ldots, x_i^r)$ for $i \in \mathcal{I}_n$. The $(r \times n)$-matrix $X^\top$ is the *transpose* of an $(n \times r)$-matrix X, i.e. $(X^\top)_{\cdot j} = X_{j\cdot}$ for all $j \in \mathcal{I}_n$.

Let $x^1, \ldots, x^r$ be n-vectors in $\mathbf{R}^n$. Then we can define an *affine combination* of these vectors as follows.

Definition 2.1.1 *Let $x^1, \ldots, x^r$ be a finite number of points in $\mathbf{R}^n$. An* affine combination *of these points is a point $x \in \mathbf{R}^n$ such that*

$$x = \sum_{i=1}^{r} \alpha_i x^i \text{ where } \sum_{i=1}^{r} \alpha_i = 1 \text{ and } \alpha_i \in \mathbf{R} \text{ for } i \in \mathcal{I}_r.$$

The set of all affine combinations of $x^1, \ldots, x^r$ is the *affine hull* of the set $\{x^1, \ldots, x^r\}$ and is denoted by $\mathrm{aff}(\{x^1, \ldots, x^r\})$. In general, for any subset $\mathcal{C}$ of $\mathbf{R}^n$, we can define the affine hull of $\mathcal{C}$, denoted $\mathrm{aff}(\mathcal{C})$, as the set of all affine combinations of any finite number of points in $\mathcal{C}$.

We define the *boundary*, *closure*, and *interior* of a set $\mathcal{C}$ in $\mathbf{R}^n$ with respect to $\mathrm{aff}(\mathcal{C})$ and denote them by $\mathrm{bd}(\mathcal{C})$, $\mathrm{cl}(\mathcal{C})$, and $\mathrm{int}(\mathcal{C})$ respectively. Also *boundedness*,

closedness, and *openness* of a subset $\mathcal{C}$ of $\mathbf{R}^n$ is defined in the usual way with respect to aff($\mathcal{C}$). A set $\mathcal{C}$ in $\mathbf{R}^n$ which is bounded and closed is *compact.*

The subset of $\mathbf{R}^n$ containing the n-vectors x such that $x_i \geq 0$ for all $i \in \mathcal{I}_n$ is denoted by $\mathbf{R}^n_+$ while the subset of $\mathbf{R}^n$ containing the n-vectors x such that $x_i > 0$ for all $i \in \mathcal{I}_n$ is denoted by $\mathbf{R}^n_{++}$. The subset $\{x \mid c^\top x \leq d\}$ of $\mathbf{R}^n$ for some nonzero vector $c \in \mathbf{R}^n$ and some number $d \in \mathbf{R}$ is called an *affine half-space* in $\mathbf{R}^n$. A *linear half-space* in $\mathbf{R}^n$ is defined by the points $\{x \mid c^\top x \leq 0\}$ for some nonzero vector $c \in \mathbf{R}^n$. The points x in $\mathbf{R}^n$ such that $c^\top x = d$ for some nonzero vector $c \in \mathbf{R}^n$ and some number $d \in \mathbf{R}$, i.e. the subset $\{x \mid c^\top x = d\}$ of $\mathbf{R}^n$, is called a *hyperplane* in $\mathbf{R}^n$.

Let $x^1, \ldots, x^r$ again be n-vectors in $\mathbf{R}^n$. Then, similar to the definition of an affine combination, one can define a *convex combination* of $x^1, \ldots, x^r$ as follows.

Definition 2.1.2 *Let $x^1, \ldots, x^r$ be a finite number of points in $\mathbf{R}^n$. A* convex combination *of these points is a point $x \in \mathbf{R}^n$ such that*

$$x = \sum_{i=1}^{r} \alpha_i x^i \text{ where } \sum_{i=1}^{r} \alpha_i = 1 \text{ and } \alpha_1, \ldots, \alpha_r \geq 0.$$

The set of all convex combinations of $x^1, \ldots, x^r$ is the *convex hull* of the set $\{x^1, \ldots, x^r\}$ and is denoted by co($\{x^1, \ldots, x^r\}$). In general, for any subset $\mathcal{C}$ of $\mathbf{R}^n$, one can define the convex hull of a set $\mathcal{C}$, denoted co($\mathcal{C}$), as the set of all convex combinations of any finite number of points in $\mathcal{C}$. A subset $\mathcal{C}$ of $\mathbf{R}^n$ is said to be *convex* if the convex hull of any two points in $\mathcal{C}$ also lies in $\mathcal{C}$. Convex sets have the property that they can be separated from points outside the set by a hyperplane.

Theorem 2.1.1 (Separating Hyperplane Theorem) *Let $\mathcal{C}$ be a nonempty convex subset of $\mathbf{R}^n$ and let $b \in \mathbf{R}^n \setminus \mathcal{C}$. Then there exists a hyperplane $\{x \mid c^\top x = d\}$ for some $c \in \mathbf{R}^n \setminus \{0\}$ and some number $d \in \mathbf{R}$ such that $c^\top x \leq d$ for all $x \in \mathcal{C}$ and $c^\top b \geq d$.*

Proof: See Murty (1988), p.524. □

Corollary 2.1.1 *Let $\mathcal{C}$ be a convex subset of $\mathbf{R}^n$ and let b be a point on the boundary of $\mathcal{C}$. Then there exists a $c \in \mathbf{R}^n \setminus \{0\}$ such that $c^\top x \leq c^\top b$ for all $x \in \mathcal{C}$.*

Proof: Follows immediately from Theorem 2.1.1. □

The hyperplane $\{x \mid c^T x = c^T b\}$ in Corollary 2.1.1 is known as the *supporting hyperplane* of C in b and it maximizes $c^T x$ on C, i.e. $\max\{c^T x \mid x \in C\} = c^T b$.

2.2 Polyhedrons, polytopes, and cones

The convex sets in this monograph mainly take the form of some polyhedron. A *polyhedron* is defined as the intersection of finitely many affine half-spaces. If a polyhedron is bounded then it is called a *polytope* while if the polyhedron is the intersection of finitely many linear half-spaces then this type of polyhedron is called a *polyhedral cone*. A polyhedron is always convex and closed.

Let A be an $(m \times n)$-matrix and b some m-vector. If $\{x \mid Ax \leq b\}$ is nonempty then this set defines a polyhedron in $\mathbf{R}^n$ or even a polytope in case this set is also bounded. A set $\{x \mid Ax \leq 0\}$ defines a polyhedral cone. A cone generated by finitely many vectors $a^1, \ldots, a^n$, denoted $\text{cone}(\{a^1, \ldots, a^n\})$, is called *a finitely generated cone* and is defined as

$$\text{cone}(\{a^1, \ldots, a^n\}) = \{\lambda_1 a^1 + \cdots + \lambda_n a^n \mid \lambda_1, \ldots, \lambda_n \geq 0\}.$$

In Schrijver (1986) p.87 it is shown that for a cone the concepts of "polyhedral" and "finitely generated" are equivalent.

Let $a^1, \ldots, a^n$ denote the column vectors of an $(m \times n)$-matrix A. Then, for any m-vector b, either $b \in \text{cone}(\{a^1, \ldots, a^n\})$ and $Ax = b$ has a nonnegative solution, or $b \notin \text{cone}(\{a^1, \ldots, a^n\})$ and there exists a hyperplane $\{x \mid y^T x = d\}$ for some n-vector $y \neq 0$ and some number d separating b strictly from $\text{cone}(\{a^1, \ldots, a^n\})$. This result is referred to in the literature as Farkas' lemma.

Lemma 2.2.1 (Farkas' lemma) *Let A be an $(m \times n)$-matrix and let b be an m-vector. Then there exists an n-vector $x \geq 0$ with $Ax = b$ if and only if there does not exist an m-vector y such that $y^T A \leq 0$ and $y^T b > 0$.*

Proof: See Schrijver (1986) □

In the literature there are a number of variants on Farkas' Lemma, among which the following one, yielding conditions on polyhedrons to be nonempty.

Corollary 2.2.1 *Let A be an $(m \times n)$-matrix and let b be an m-vector. Then the system $Ax \leq b$ has a solution $x \geq 0$ if and only if there does not exist a vector $y \leq 0$ with $y^\top b > 0$ and $y^\top A \leq 0$.*

Proof: Let A' be the $(m \times (m+n))$-matrix $(I \ A)$ where I denotes the $(m \times m)$-matrix with columns $e(1), \ldots, e(m)$. Then $Ax \leq b$ has a solution $x \geq 0$ if and only if $A'x' = b$ has a solution $x' \geq 0$. Application of Lemma 2.2.1 to this system yields the required result. □

Let $\mathcal{M} = \{x \mid Ax \leq b\}$ define a nonempty polyhedron in $\mathbf{R}^n$. Then we call $\mathcal{F}$ a *face* of $\mathcal{M}$ if $\mathcal{F}$ is either the empty set, or $\mathcal{M}$ itself, or the set $\mathcal{H} \cap \mathcal{M}$ for some supporting hyperplane $\mathcal{H}$ of $\mathcal{M}$. If $\mathcal{F}$ consists of a point then the face $\mathcal{F}$ is called a *vertex* of $\mathcal{M}$. If $\mathcal{F}$ consists of a line segment or a half line then $\mathcal{F}$ is called an *edge* of $\mathcal{M}$, while a face $\mathcal{F}$ of dimension one lower than the dimension of $\mathcal{M}$ is called a *facet* of $\mathcal{M}$.

Suppose $\mathcal{H} = \{x \mid c^\top x = d\}$ defines a supporting hyperplane of a polyhedron $\mathcal{M}$ in $\mathbf{R}^n$ for some $c \in \mathbf{R}^n \setminus \{0\}$ and some number d. Let $\mathcal{F}$ be the face of $\mathcal{M}$ determined by $\mathcal{H} \cap \mathcal{M}$. Corollary 2.1.1 implies that the set $\mathcal{H} \cap \mathcal{M}$ is determined by maximizing the linear function $c^\top x$ on $\mathcal{M}$. Hence $\mathcal{F}$ is the set of points in $\mathcal{M}$ attaining $\max\{c^\top x \mid x \in \mathcal{M}\}$. The next theorem shows that the solution to this linear programming problem can also be found by solving an alternative one. In linear programming theory this theorem is referred to as the Duality Theorem of Linear Programming.

Theorem 2.2.1 (Duality Theorem of Linear Programming) *Let A be an $(m \times n)$-matrix, b an m-vector, and c an n-vector. Then*

$$\max\{c^\top x \mid Ax \leq b\} = \min\{y^\top b \mid y \geq 0, \ y^\top A = c^\top\} \tag{2.1}$$

provided that both sets in (2.1) *are nonempty.*

Proof: See Schrijver (1986) p.90. □

If one considers the maximization problem in (2.1) as the *primal* problem then the minimization problem in (2.1) is referred to as the *dual* of the maximization problem and conversely.

Let x^0 be a solution to $\max\{c^\top x \mid Ax \leq b\}$ and let y^0 be a solution to its dual $\min\{y^\top b \mid y \geq 0,\ y^\top A = c^\top\}$. Then (2.1) says that $c^\top x^0 = y^{0\top} b$. Because $c^\top = y^{0\top} A$ it follows that

$$y^{0\top}(b - Ax^0) = 0. \tag{2.2}$$

This implies that if a component of y^0 is positive then the corresponding inequality in $Ax \leq b$ is satisfied in x^0 with equality. In linear programming theory this phenomenon is known as the *complementarity property*. Also, it follows that

$$x^{0\top}(c - A^\top y^0) = 0. \tag{2.3}$$

To solve the optimization problems in (2.1) one may apply the simplex method as introduced in Dantzig (1951) to the primal. An alternative way is to find $x \in \mathbf{R}^n$ and $y \in \mathbf{R}^m$ such that

$$\begin{aligned} &y^\top A = c^\top \\ &Ax \leq b \\ &y^\top(b - Ax) = 0 \\ &y \geq 0. \end{aligned}$$

These conditions are known as the *Karush-Kuhn-Tucker conditions* (see Karush (1939), and Kuhn and Tucker (1951)).

Corollary 2.2.2 *Let A be an $(m \times n)$-matrix and let b be an m-vector and c an n-vector. Then $x^0 \in \mathbf{R}^n$ and $y^0 \in \mathbf{R}^m$ are a solution to*

$$\begin{aligned} &y^\top A = c^\top \\ &Ax \leq b \\ &y^\top(b - Ax) = 0 \\ &y \geq 0 \end{aligned} \tag{2.4}$$

if and only if x^0 solves $\max\{c^\top x \mid Ax \leq b\}$ and y^0 solves $\min\{y^\top b \mid y \geq 0,\ y^\top A = c^\top\}$.

Proof: Let $x^0 \in \mathbf{R}^n$ and $y^0 \in \mathbf{R}^m$ be a solution to the Karush-Kuhn-Tucker conditions in (2.4). Suppose there exists an $\hat{x} \in \mathbf{R}^n$ such that $A\hat{x} \leq b$ and $c^\top \hat{x} > c^\top x^0$. As $c^\top = y^{0\top} A$ it holds that $y^{0\top}(A\hat{x} - b) > 0$. Since $y^0 \geq 0$ this implies that there exists some $j \in \mathcal{I}_m$ such that $y_j^0 > 0$ and $A_{j\cdot}\hat{x} > b_j$. This contradicts $A\hat{x} \leq b$. Hence x^0 solves $\max\{c^\top x \mid Ax \leq b\}$. By the Duality Theorem of Linear Programming it

follows that y^0 solves $\min\{y^\top b \mid y \geq 0,\ y^\top A = c^\top\}$. The reverse follows immediately from the discussion above. □

Let $\mathcal{M} = \{x \mid Ax \leq b\}$ be a nonempty polyhedron where A is an $(m \times n)$-matrix and b an m-vector. For $\mathcal{G} \subset \mathcal{I}_m$ the set $\mathcal{F}(\mathcal{G})$ defined by

$$\mathcal{F}(\mathcal{G}) = \{x \in \mathcal{M} \mid a_i^\top x = b_i \text{ for all } i \in \mathcal{G}\} \tag{2.5}$$

is a face of $\mathcal{M}$. In case the system $Ax \leq b$ contains constraints which are implied by other constraints then $\mathcal{M}$ contains so-called *redundant* constraints. Redundant constraints can be removed although deleting one constraint can make another redundant constraint irredundant. The possible occurence of redundant constraints in $Ax \leq b$ may cause the set of integers $\mathcal{G}$ defining a nonempty face $\mathcal{F}$ such that $\mathcal{F}(\mathcal{G}) = \mathcal{F}$ to be nonunique. Throughout this monograph we will however assume each constraint to be irredundant.

Furthermore we assume in this monograph that the polyhedron $\mathcal{M}$ has dimension n and is *simple*. This means that every nonempty face $\mathcal{F}(\mathcal{G})$ of $\mathcal{M}$ in $\mathbf{R}^n$ as defined in (2.3) has dimension equal to $n - |\mathcal{G}|$, where $|\mathcal{G}|$ denotes the number of elements in the set $\mathcal{G}$. Notice that simpleness is different from redundancy. The polytope defined by $\mathcal{M} = \{x \in \mathbf{R}^3 \mid -x_1 + x_3 \leq 0,\ -x_2 + x_3 \leq 0,\ x_2 + x_3 \leq 1,\ x_1 + x_3 \leq 1,\ -x_3 \leq 0\}$ is not simple as the face $\mathcal{F}(\mathcal{I}_4) = \{(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})^\top\}$ has dimension zero while $n - |\mathcal{I}_4| = -1$. However, $\mathcal{M}$ does not contain any redundant constraints.

2.3 Fixed points, stationary points, and complementarity

Let $\mathcal{C}$ be a subset of $\mathbf{R}^n$ and f a function from $\mathcal{C}$ to some subset of $\mathbf{R}^n$. A point $x^* \in \mathcal{C}$ such that $f(x^*) = x^*$ is called a *fixed point* of f. In 1910 the Dutch mathematician L.E.J. Brouwer proved the following theorem.

Theorem 2.3.1 *Let $\mathcal{C}$ be a nonempty, compact, and convex subset of $\mathbf{R}^n$. Then any continuous function $f : \mathcal{C} \to \mathcal{C}$ has a point $x^* \in \mathcal{C}$ such that $f(x^*) = x^*$.*

Proof: See Brouwer (1912). □

Let $\mathcal{X}$ be some set of nonempty subsets of $\mathbf{R}^n$. Then a function $F : \mathcal{C} \to \mathcal{X}$ is called a *point-to-set mapping* from $\mathcal{C}$ to $\mathbf{R}^n$. We call F *upper semicontinuous at a point* $\hat{x} \in \mathcal{C}$ if for every open set $\mathcal{B}$ containing $F(\hat{x})$ there exists a neighbourhood $\mathcal{N}(\hat{x})$ of $\hat{x}$ such that $F(x)$ is contained in $\mathcal{B}$ for every $x \in \mathcal{N}(\hat{x})$. F is *upper semicontinuous on* $\mathcal{C}$ if $F(x)$ is compact and upper semicontinuous at every $x \in \mathcal{C}$.

In 1941 Kakutani extended Brouwer's result to point-to-set mappings.

Theorem 2.3.2 *Let $\mathcal{C}$ be a nonempty, compact, convex subset of $\mathbf{R}^n$ and $\mathcal{X}$ the set of nonempty convex subsets of $\mathcal{C}$. Then any upper semicontinuous point-to-set mapping $F : \mathcal{C} \to \mathcal{X}$ has a point $x^* \in \mathcal{C}$ such that $x^* \in F(x^*)$.*

Proof: See Kakutani (1941). □

A point x^* in $\mathbf{R}^n$ is called a *stationary point* of a function $f : \mathcal{C} \to \mathbf{R}^n$ on $\mathcal{C}$ if $x^* \in \mathcal{C}$ and $x^{*\mathsf{T}} f(x^*) \geq x^{\mathsf{T}} f(x^*)$ for all $x \in \mathcal{C}$. The stationary point problem (*SPP*) of f on $\mathcal{C}$ is to find a stationary point of f on $\mathcal{C}$. The point x^* being a stationary point of f on $\mathcal{C}$ is equivalent to x^* solving the optimization problem given by $\max\{x^{\mathsf{T}} f(x^*) \mid x \in \mathcal{C}\}$. In case $\mathcal{C}$ is a polyhedron this optimization problem is a linear programming problem as it maximizes the linear function $x^{\mathsf{T}} f(x^*)$ over the set $\mathcal{C}$. Furthermore it can easily be seen that the set $\{x \mid x^{\mathsf{T}} f(x^*) = x^{*\mathsf{T}} f(x^*)\}$ is a supporting hyperplane of $\mathcal{C}$. Figure 2.3.1 illustrates these properties of a stationary point.

Theorem 2.3.1 implies the following result.

Corollary 2.3.1 *Let $\mathcal{C}$ be a nonempty, compact, convex set in $\mathbf{R}^n$ and $f : \mathcal{C} \to \mathbf{R}^n$ a continuous function. Then there exists a stationary point x^* of f on $\mathcal{C}$.*

Let $f : \mathbf{R}^n_+ \to \mathbf{R}^n$ be a function. The problem to find a point x^* such that

$$x^{*\mathsf{T}} f(x^*) = 0,\ f(x^*) \geq 0,\ x^* \geq 0$$

is called the *nonlinear complementarity problem* (*NLCP*) as introduced in Cottle (1966). In case f is affine, i.e. $f(x) = Mx + q$ for some given $n \times n$-matrix M and some given n-vector q, then the well-known *linear complementarity problem* is obtained. The following theorem establishes the relationship between the nonlinear complementarity problem and a stationary point problem.

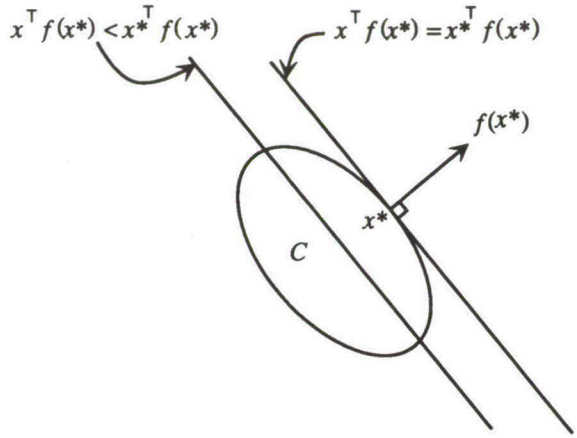


FIGURE 2.3.1: The point x^* is a stationary point of f on $\mathcal{C}$. Then x^* maximizes $x^T f(x^*)$ over $\mathcal{C}$. The hyperplane $\{x \mid x^T f(x^*) = x^{*T} f(x^*)\}$ is a supporting hyperplane of $\mathcal{C}$ in x^*.

Theorem 2.3.3 *Let $f : \mathbf{R}^n_+ \to \mathbf{R}^n$ be a function. Then a point $x^* \in \mathbf{R}^n_+$ solves the NLCP to f if and only if x^* is a stationary point of $-f$ on $\mathbf{R}^n_+$.*

Proof: Let x^* be a solution to the *NLCP* but assume that x^* is not a stationary point of $-f$ on $\mathbf{R}^n_+$. Then there exists a point $\hat{x} \in \mathbf{R}^n_+$ such that $-\hat{x}^T f(x^*) > -x^{*T} f(x^*)$. Because x^* solves the *NLCP* this result reduces to $\hat{x}^T f(x^*) < 0$ which is in contradiction to the fact that $\hat{x} \in \mathbf{R}^n_+$ and $f(x^*) \geq 0$ implies that $\hat{x}^T f(x^*) \geq 0$.

Let x^* be a stationary point of $-f$ on $\mathbf{R}^n_+$. Then by definition it holds that $-x^T f(x^*) \leq -x^{*T} f(x^*)$ for all $x \in \mathbf{R}^n_+$. For $x = 0$ this implies that $0 \leq -x^{*T} f(x^*)$. For $x = x^* + e(i)$ this implies that $f_i(x^*) \geq 0$ for all $i \in \mathcal{I}_n$. Then $-x^{*T} f(x^*) \leq 0$. Hence $x^{*T} f(x^*) = 0$. □

Part B of this monograph will be entirely dedicated to the computation of solutions to complementarity problems. The algorithms introduced in that part are all based on the notion of equivalence between stationary point problems and complementarity problems.

Chapter 3

Path-following algorithms

The algorithms presented in this monograph are so-called path-following algorithms. They generate a path of points in some convex set starting in a point contained in this set and end up with an approximating solution to the problem unless they diverge towards infinity. This path is piecewise linear, i.e. the path consists of a sequence of adjacent line segments. Two line segments are adjacent if they share a common end point. An end point of a line segment on the path generated by the algorithm is therefore either the starting point of the path, or an end point of an adjacent line segment, or an approximating solution to the problem.

A sequence of adjacent line segments can be generated by moving subsequently from one end point of a line segment in the sequence to its other end point. Every line segment can be seen as corresponding to a set of solutions to a system of linear equations with one degree of freedom. On the variables of this system certain restrictions hold. In order to follow the set of solutions we can vary just one of the variables in the system of equations while maintaining the validity of the restrictions. An end point of a line segment can then be seen as corresponding to a solution to the system of linear equations where one of the constraints on the variables is binding. If one makes this constraint unbinding then eventually some other constraint may become binding, resulting in the other end point of the line segment. This procedure can be carried out by applying a technique known from linear programming as the pivot procedure. Section 1 of this chapter describes this pivot procedure in detail.

In order to apply the linear programming pivoting technique one should obtain a

system of linear equations from the problem under consideration. The problems in this monograph however will frequently be nonlinear. Thus linearizations are needed to obtain a system of linear equations. In this monograph problems are linearized either by taking the so-called first-order Taylor-expansion of the nonlinear terms in the obtained system or by piecewise linearizing the problem on a simplicial subdivision of the underlying set. Section 2 describes the concept of a simplicial subdivision of a set and discusses some examples of simplicial subdivisions. In Section 3 the linearizations mentioned above are described.

The points on the path generated by the algorithms all fulfil certain parametrized conditions. This means that for each point on the path there exists some value of a parameter such that the conditions are satisfied in this point. It will be these conditions which contain the main idea behind the algorithms presented in this monograph and from which the system of linear equations is derived. The parametrized conditions can be interpreted as a homotopy problem with the parameter as the underlying homotopy variable. Section 4 describes the homotopy problem in a general setting.

3.1 Generating end points of line segments

Consider the following system of equations in the variables $x \in \mathbf{R}^k$:

$$\begin{aligned} Ax &= b \\ x &\geq 0, \end{aligned} \tag{3.1}$$

where A is a given $(n \times k)$-matrix, $n < k$, and b is a given n-vector. Assume that the matrix A contains n linearly independent row vectors and that the set $\{x \mid Ax = b,\ x \geq 0\}$ is $(k-n)$-dimensional and simple. Denote the jth column vector of A by a^j and let $\mathcal{T}$ be some subset of $\mathcal{I}_k$ such that $\mid \mathcal{T} \mid = n$ and a^j $(j \in \mathcal{T})$ are linearly independent. Then a^j $(j \in \mathcal{T})$ form a basis of $\mathbf{R}^n$, i.e. there exists a unique n-vector $\bar{x}$ such that $\bar{x}_j = 0$ for all $j \notin \mathcal{T}$ and $\bar{x}_j$ $(j \in \mathcal{T})$ solve the system of equations

$$\sum_{j \in \mathcal{T}} a^j x_j = b \tag{3.2}$$

In this system x_j $(j \in \mathcal{T})$ are called the *basic variables* and x_j $(j \notin \mathcal{T})$ are called the *nonbasic variables*. Let B be the matrix consisting of all the columns of A associated

with the basic variables, called the *basic columns*, i.e. $B = (a^j,\ j \in \mathcal{T})$, and let x^B be the column vector with the basic variables associated with the columns of B. Then $\overline{x}$ is called a *basic solution* to (3.1) and is obtained by taking $\overline{x}_j = 0$ $(j \notin \mathcal{T})$ and $\overline{x}^B = B^{-1}b$. If $\overline{x}^B \geq 0$ then $\overline{x}$ is a *feasible basic solution* to (3.1).

Let $x^D = (x_j \mid j \notin \mathcal{T})$ be the $(k-n)$-vector containing the nonbasic variables in (3.1). Without loss of generality we assume that $x = (x^B, x^D)^\top$. Suppose that $\overline{x}$ is a feasible basic solution to (3.1) with $\overline{x}_j > 0$ $(j \in \mathcal{T})$. For any $t \notin \mathcal{T}$, such a solution $\overline{x}$ satisfies the system

$$\begin{aligned} &\textstyle\sum_{j\in\mathcal{T}} a^j x_j + a^t x_t = b \\ &x_j \geq 0\ (j \in \mathcal{T}),\ x_t \geq 0, \\ &x_j = 0\ (j \notin \mathcal{T} \cup \{t\}). \end{aligned} \tag{3.3}$$

If also $\overline{x}_t = 0$, then $\overline{x}$ is an end point of a line segment of solutions to (3.1). This line segment can be generated by increasing the nonbasic variable x_t in (3.3) from zero. When raising x_t from zero, every basic variable x_j $(j \in \mathcal{T})$ in (3.3) will be adapted in order to maintain the equality in (3.3).

As the column vectors a^j $(j \in \mathcal{T})$ constitute a basis of $\mathbf{R}^n$ there exists a unique n-vector y^B with components y_j $(j \in \mathcal{T})$ such that

$$\sum_{j\in\mathcal{T}} a^j y_j = a^t, \tag{3.4}$$

namely $y^B = B^{-1}a^t$. Subtracting x_t times equation (3.4) from equation (3.2) at the solution $\overline{x}$ results in

$$\sum_{j\in\mathcal{T}} a^j(\overline{x}_j - x_t y_j) + x_t a^t = b. \tag{3.5}$$

This equation shows that if one increases x_t from zero then the basic variable x_j $(j \in \mathcal{T})$ in (3.3) is changed from $\overline{x}_j$ with rate y_j in order to maintain the equalities in (3.1).

Suppose there exists some $i \in \mathcal{T}$ such that $y_i > 0$. Then there is some positive value of x_t such that for some $h \in \mathcal{T}$ the basic variable x_h becomes zero while all other basic variables x_j $(j \in \mathcal{T} \setminus \{h\})$ are still nonnegative. This occurs when x_t is equal to $\hat{x}_t$ where

$$\hat{x}_t = \frac{\overline{x}_h}{y_h} \tag{3.6}$$

for the basic variable x_h such that

$$\frac{\overline{x}_h}{y_h} = \min_{\{j|y_j>0\}} \left\{ \frac{\overline{x}_j}{y_j} \right\}. \tag{3.7}$$

Assume for the moment that the index h in (3.7) is unique. Let $\hat{x}_j = \overline{x}_j - \hat{x}_t y_j$ ($j \in \mathcal{T} \setminus \{h\}$), $\hat{x}_h = 0$, and $\hat{x}_j = 0$ ($j \notin \mathcal{T} \cup \{t\}$). Then (3.3) determines a line segment of solutions to (3.1) with end points $\overline{x}$ and $\hat{x}$, given by

$$\{\overline{x} - \lambda y \mid 0 \leq \lambda \leq \hat{x}_t\}, \tag{3.8}$$

where y_j ($j \in \mathcal{T}$) are defined as in (3.4), $y_t = -1$, and $y_j = 0$ for all $j \notin \mathcal{T} \cup \{t\}$. The point $\hat{x}$ is again a basic feasible solution to (3.1).

Suppose there does not exist some $i \in \mathcal{T}$ such that $y_i > 0$, giving $y \leq 0$. Then in (3.5) x_t can be raised towards infinity without x_h becoming zero for some $h \in \mathcal{T}$. No other end point of this line segment exists. In fact (3.5) determines a *half line* or *ray* of solutions to (3.1) with $\overline{x}$ as its unique end point, given by

$$\{\overline{x} - \lambda y \mid 0 \leq \lambda\}. \tag{3.9}$$

Notice that in (3.9) the vector $\overline{x}$ is a basic feasible solution to (3.1) and the vector y is a solution to the homogeneous system of equations $Ay = 0$.

The procedure to generate a line segment of solutions to (3.1) is known as the *pivot procedure*. The pivot column a^t in (3.3) is brought into system (3.3) by raising the corresponding *entering variable* x_t from zero. The column a^h leaving the basis is determined by applying the *minimum-ratio test* in (3.7). The variable x_h becoming zero when applying the pivot procedure is called the *leaving variable*. Finally the column a^t is pivoted into (the inverse of) the basis matrix B to replace a^h.

Until now we have assumed that the minimum-ratio test in (3.7) results in a unique leaving variable x_h and a positive minimum-ratio. Suppose that this assumption does not hold, i.e. the minimum-ratio in (3.6) is zero or the minimum-ratio test does not result in a unique leaving variable according to (3.7).

In case the minimum-ratio in (3.6) is zero then $\hat{x}_t = 0$ and the line segment in (3.8) reduces to a single point $\overline{x}$. Pivoting the column a^t into the basis then results in a new basis where x_t replaces some basic variable x_h but generates the same end point $\overline{x} = \hat{x}$. The end point $\overline{x}$ is therefore called *degenerate*.

If the minimum-ratio test in (3.7) gives rise to more than one possible leaving variable then we can make a choice among these possibilities. Let $\mathcal{H}$ be a set of integers such that (3.7) holds for all $h \in \mathcal{H}$ in the end point $\bar{x}$. Choose some $h_1 \in \mathcal{H}$ as the leaving variable. Notice that for all $h \in \mathcal{H}$ it holds that $\bar{x}_h - \hat{x}_t y_h = 0$. Pivoting a^t into the basis therefore leads to an end point $\hat{x}$ with $\hat{x}_h = 0$ for all $h \in \mathcal{H}$. The basis in $\hat{x}$ therefore contains basic variables equal to zero, namely x_h $(h \in \mathcal{H} \setminus \{h_1\})$. In accordance with the results from the previous paragraph one can see that $\hat{x}$ might be a degenerate end point of the line segment.

Since degeneracy can easily be avoided by taking lexicographic pivoting or perturbing the right hand side of the system of equations, we assume in this monograph that degeneracy during any pivot procedure will not occur.

3.2 Simplicial subdivisions

This section describes a subdivision of a convex set in $\mathbf{R}^n$ into so-called simplices. Several of the algorithms introduced in this monograph apply some simplicial subdivision of the set on which the problem to be solved is defined. This in order to obtain a linear system. The use of different kinds of simplicial subdivisions to underly a particular algorithm may have large consequences for the efficiency of the algorithm as well as for the accuracy of the approximation of a solution by the algorithm, see e.g. Todd (1976), Doup (1988), and Dang (1991b).

Let $\mathcal{C}$ be some m-dimensional convex set in $\mathbf{R}^n$. A *simplicial subdivision* of $\mathcal{C}$ subdivides $\mathcal{C}$ into so-called m-simplices. An *m-simplex* or *m-dimensional* simplex is the convex hull of $m+1$ affinely independent points. The points $y^1, \ldots, y^{t+1}$ in $\mathbf{R}^n$ are called *affinely independent* if

$$\sum_{i=1}^{t+1} \lambda_i y^i = 0 \text{ and } \sum_{i=1}^{t+1} \lambda_i = 0 \text{ imply } \lambda_i = 0,\ i = 1, \ldots, t+1.$$

A simplex is defined as follows.

Definition 3.2.1 *Let $y^1, \ldots, y^{t+1}$ be $t+1$ affinely independent points in $\mathbf{R}^n$. Then the* t-dimensional simplex *or* t-simplex *$\sigma(y^1, \ldots, y^{t+1})$ is the convex hull of $y^1, \ldots, y^{t+1}$, i.e.*

$$\sigma(y^1, \ldots, y^{t+1}) = \text{co}(\{y^1, \ldots, y^{t+1}\}).$$

The points $y^1, \ldots, y^{t+1}$ are called the *vertices* of the simplex $\sigma = \sigma(y^1, \ldots, y^{t+1})$. A *face* of a t-simplex σ is a k-simplex τ $(k \leq t)$ whose vertices are also vertices of σ. If $k = t - 1$ we call τ a *facet* of σ. If y is the vertex of σ which is not a vertex of a facet τ of σ we say that τ is the facet of σ *opposite* the vertex y.

A collection of simplices is said to be a *simplicial subdivision* of C if it fulfils the properties given in Definition 3.2.2.

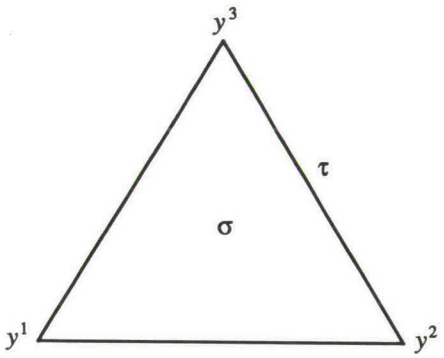


FIGURE 3.2.1: The 2-simplex σ with vertices y^1, y^2, y^3 has faces $\emptyset$, $\{y^1\}$, $\{y^2\}$, $\{y^3\}$, $\text{co}(\{y^1, y^2\})$, $\text{co}(\{y^1, y^3\})$, $\text{co}(\{y^2, y^3\})$, and $\text{co}(\{y^1, y^2, y^3\})$. $\text{co}(\{y^1, y^2\})$, $\text{co}(\{y^1, y^3\})$, and $\text{co}(\{y^2, y^3\})$ are the facets of σ. The facet $\tau = \text{co}(\{y^2, y^3\})$ is the facet of σ opposite the vertex y^1.

Definition 3.2.2 *Let C be an m-dimensional convex set in $\mathbf{R}^n$. A collection $\mathcal{G}$ of m-simplices is a* simplicial subdivision *of C if*

i) C is the union of all simplices in $\mathcal{G}$;

ii) the intersection of two simplices in $\mathcal{G}$ is either empty or a common face;

iii) each facet either belongs to the boundary of C and is a facet of just one simplex of

$\mathcal{G}$ or does not belong to the boundary of C and is a facet of exactly two simplices of $\mathcal{G}$.

Let $C = [0,1]$ and let $\mathcal{G}^1$ be the collection of simplices $\sigma(y^1, y^2)$ in $\mathbf{R}$ such that $y^1 = 2^{-(k+1)}$ and $y^2 = 2^{-k}$, $k = 0, 1, \ldots$. Then $\mathcal{G}^1$ is not a simplicial subdivision of $[0,1]$ as $\cup_{\sigma \in \mathcal{G}^1} \sigma \neq C$, thereby violating the first condition in Definition 3.2.2. Figure 3.2.2 illustrates the other two conditions in Definition 3.2.2.

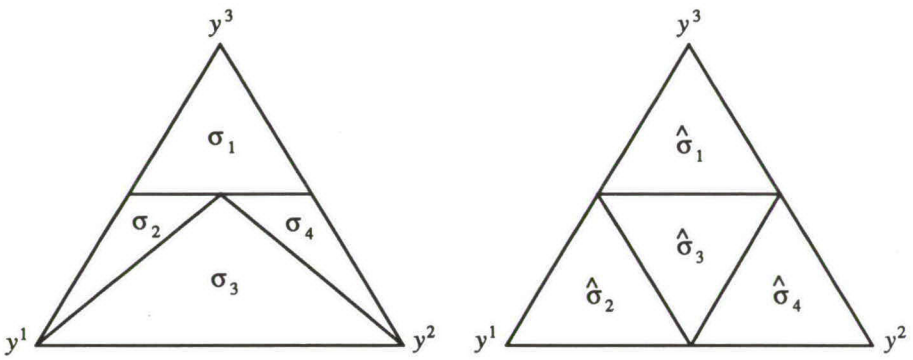


FIGURE 3.2.2: $C = \text{co}(\{y_1, y_2, y_3\})$. Then the collection $\{\sigma_1, \sigma_2, \sigma_3, \sigma_4\}$ is not a simplicial subdivision of C but $\{\hat{\sigma}_1, \hat{\sigma}_2, \hat{\sigma}_3, \hat{\sigma}_4\}$ is.

Two different simplices sharing a common facet or one simplex being a facet of the other one are called *adjacent*. The properties defining a simplicial subdivision $\mathcal{G}$ of an m-dimensional convex set C imply the existence of sequences of adjacent m-simplices in $\mathcal{G}$ as is illustrated in Figure 3.2.3. Suppose σ^1 and σ^2 are two adjacent simplices in such a sequence sharing a common facet τ. Then σ^2 can be obtained from σ^1 by replacing the vertex of σ^1 opposite τ, say y, by the vertex of σ^2 opposite τ, say $\overline{y}$. The movement from an m-simplex σ^1 to an adjacent m-simplex σ^2 is called a *replacement step*. Figure 3.2.4 illustrates a replacement step between σ^1 and σ^2 in the two-dimensional case.

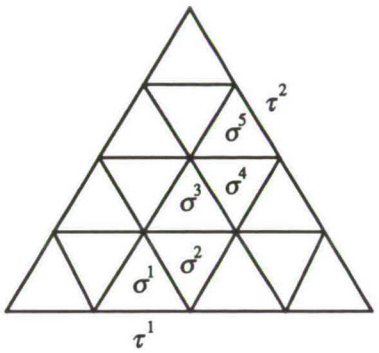


FIGURE 3.2.3: The simplices τ^1 and σ^1, σ^5 and τ^2, as well as σ^j and σ^{j+1}, $j = 1, \ldots, 4$, are adjacent. The sequence $(\tau^1, \sigma^1, \sigma^2, \ldots, \sigma^5, \tau^2)$ gives an example of a sequence of adjacent simplices in $\mathcal{G}$.

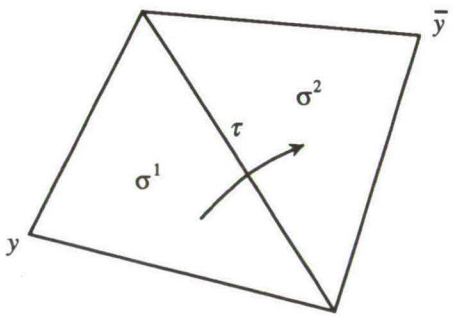


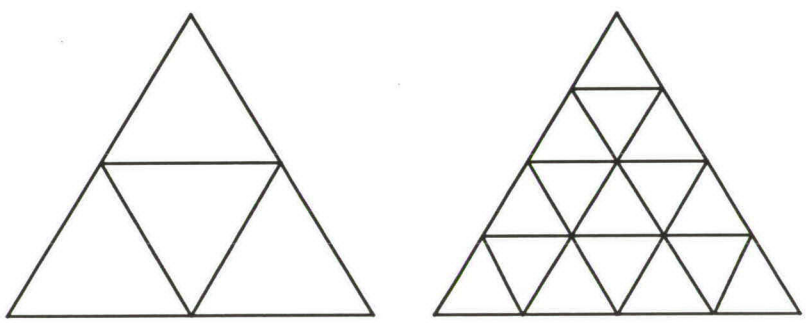
FIGURE 3.2.4: Replacement step between σ^1 and σ^2 in the two-dimensional case where the vertex $\bar{y}$ of σ^2 opposite τ replaces the vertex y of σ^1 opposite τ.

By subsequently making a replacement step in each of the simplices one generates a sequence of adjacent m-simplices in a simplicial subdivision $\mathcal{G}$. Such a sequence takes the form of a *cycle* if the replacement procedure returns to the simplex with which the sequence was started.

To compare simplicial subdivisions of a convex set $\mathcal{C}$ we define the *mesh* of a simplicial subdivision $\mathcal{G}$ of $\mathcal{C}$ by

$$\text{mesh } \mathcal{G} = \sup_{\sigma \in \mathcal{G}} \{\text{diam } \sigma\}$$

where diam $\sigma = \max\{\|x-y\| \mid x, y \in \sigma\}$ denotes the *diameter* of the simplex σ. Given the definition of an m-simplex it may be clear that the diameter of an m-simplex is the largest distance between two vertices of this simplex.



(a): Simplicial subdivision $\mathcal{G}$ (b) Simplicial subdivision $\hat{\mathcal{G}}$

FIGURE 3.2.5: The simplicial subdivision $\mathcal{G}$ has a mesh being two times larger than the mesh of the simplicial subdivision $\hat{\mathcal{G}}$.

The mesh of a simplicial subdivision measures its refinement. A simplicial subdivision is said to become *finer* if its mesh becomes smaller. In Figure 3.2.5, $\hat{\mathcal{G}}$ is a finer simplicial subdivision than $\mathcal{G}$.

A simplicial subdivision is also called a *triangulation* refering to the two-dimensional case. In this monograph we will use the term simplicial subdivision instead of triangulation. An exception is made when refering to specific simplicial subdivisions whose names incorporate the term triangulation. The next four examples briefly describe well-known simplicial subdivisions in the literature on this subject. For an extensive review we refer the reader to the appropriate literature. Let S^n be the n-dimensional unit simplex, i.e. $S^n = \{x \in \mathbf{R}_+^{n+1} \mid \sum_{j=1}^{n+1} x_j = 1\}$.

Q-triangulation of S^n. The Q-triangulation is probably the most well-known simplicial subdivision. It was proposed in Freudenthal (1942) and it is defined as follows.

Definition 3.2.3 (Q-triangulation) *The* Q-triangulation *of* S^n *with grid size* m^{-1}, $m \in \mathbf{N}$, *is the collection* $\mathcal{Q}$ *of* n*-simplices* $\sigma(y^1, \pi)$ *with vertices* $y^1, \ldots, y^{n+1}$ *in* S^n *such that*

i) each component of y^1 *is a nonnegative multiple of* m^{-1};

ii) $\pi = (\pi_1, \ldots, \pi_n)$ *is a permutation of the elements in* $\mathcal{I}_n$;

iii) $y^{i+1} = y^i + m^{-1} q(\pi_i)$, $i = 1, \ldots, n$,

where $q(j) = e(j) - e(j+1)$, $j = 1, \ldots, n$.

An example of the Q-triangulation of S^n is given for the two-dimensional case with grid sizes $\frac{1}{2}$ and $\frac{1}{4}$ in Figure 3.2.5.

Let $\sigma(y^1, \pi)$ with vertices $y^1, \ldots, y^{n+1}$ and $\overline{\sigma}(\overline{y}^1, \overline{\pi})$ be two adjacent n-simplices in $\mathcal{Q}$ with a common facet τ. Let y^p denote the vertex of σ opposite the common facet τ. Then Table 3.2.1 summarizes the consequences of making a replacement step with the vertex y^p of σ on the parameters y^1 and π of σ.

V-triangulation of S^n. The V-triangulation of the unit simplex S^n was first introduced in Doup and Talman (1987). This simplicial subdivision appears to be very useful when it underlies simplicial algorithms incorporating the possibility of choosing an arbitrary starting point. Applying the Q-triangulation for this purpose limits the choice of possible starting points in S^n.

Let v be an arbitrarily chosen point in S^n. Define the projection vector $p(\mathcal{I})$ of v on the face $S^n(\mathcal{I}) := \{p \in S^n \mid p_i = 0,\ i \notin \mathcal{I}\}$ of S^n, $\mathcal{I} \subset \mathcal{I}_{n+1}$, as follows.

TABLE 3.2.1: Replacement step in $\mathcal{Q}$.

	$\overline{y}^1$	$\overline{\pi}$
$p = 1$	$y^1 + m^{-1}q(\pi_1)$	$(\pi_2, \ldots, \pi_n, \pi_1)$
$1 < p < n+1$	y^1	$(\pi_1, \ldots, \pi_{p-2}, \pi_p, \pi_{p-1}, \pi_{p+1}, \ldots, \pi_n)$
$p = n+1$	$y^1 - m^{-1}q(\pi_n)$	$(\pi_n, \pi_1, \ldots, \pi_{n-1})$

Definition 3.2.4 *Let $\mathcal{I}$ be a nonempty subset of $\mathcal{I}_{n+1}$ and let $\mathcal{I}^0$ be the set given by $\mathcal{I}^0 = \{i \in \mathcal{I} \mid v_i = 0\}$. Then the projection vector $p(\mathcal{I})$ of v on $S^n(\mathcal{I})$ is given by*

$$p_h(\mathcal{I}) = \begin{cases} 0, & h \notin \mathcal{I} \\ \dfrac{v_h(1+\mid \mathcal{I}^0 \mid)}{\sum_{k\in\mathcal{I}} v_k + \mid \mathcal{I}^0 \mid}, & h \in \mathcal{I} \setminus \mathcal{I}^0 \\ \dfrac{1 - \sum_{k\in\mathcal{I}} v_k}{\sum_{k\in\mathcal{I}} v_k + \mid \mathcal{I}^0 \mid}, & h \in \mathcal{I}^0. \end{cases}$$

For $\mathcal{I} = \emptyset$ we define $p(\emptyset) = v$.

For a discussion of these and other projections we refer to Doup and Talman (1987).

Given these projections of v on the boundary of S^n, S^n is first subdivided into subsets $A(\gamma(\mathcal{T}))$ where $\gamma(\mathcal{T}) = (\gamma_1, \ldots, \gamma_t)$ is a permutation of the t elements in some subset $\mathcal{T}$ of $\mathcal{I}_{n+1}$, such that $v_i > 0$ for some $i \notin \mathcal{T}$. The set $A(\gamma(\mathcal{T}))$ is defined as

$$A(\gamma(\mathcal{T})) = \{p \in S^n \mid p = v + \textstyle\sum_{h=1}^{t} \alpha(\gamma_h) q(\gamma_h) \text{ with } 0 \le \alpha(\gamma_t) \le \cdots \le \alpha(\gamma_1) \le 1\},$$

where $q(\gamma_h) = p(\{\gamma_1, \ldots, \gamma_h\}) - p(\{\gamma_1, \ldots, \gamma_{h-1}\})$ for $h = 1, \ldots, t$. Then $A(\gamma(\mathcal{T}))$ is subdivided into t-simplices in the following way.

Definition 3.2.5 *Let $\mathcal{T}$ be a feasible subset of $\mathcal{I}_{n+1}$ and let $\gamma(\mathcal{T}) = (\gamma_1, \ldots, \gamma_t)$ be a permutation of the elements of $\mathcal{T}$. The $V(\gamma(\mathcal{T}))$-triangulation with grid size m^{-1} of a nonempty set $A(\gamma(\mathcal{T}))$ is the collection $\mathcal{V}(\gamma(\mathcal{T}))$ of t-simplices $\sigma(y^1, \pi)$ with vertices $y^1, \ldots, y^{t+1}$ such that*

i) $y^1 = v + \sum_{k=1}^{t} d(\gamma_k) m^{-1} q(\gamma_k)$ *with integers* $d(\gamma_k)$ *satisfying* $0 \leq d(\gamma_t) \leq \cdots \leq d(\gamma_1) \leq m-1$;

ii) $\pi = (\pi_1, \ldots, \pi_t)$ *is a permutation of the elements of* $\mathcal{T}$ *such that* $p > p'$ *if* $d(\pi_{p'}) = d(\pi_p)$ *when for some* i, $1 < i \leq t$, $\pi_{p'} = \gamma_{i-1}$ *and* $\pi_p = \gamma_i$;

iii) $y^{i+1} = y^i + m^{-1} q(\pi_i)$, $i = 1, \ldots, t$.

Let the union of all sets $A(\gamma(\mathcal{T}))$ over all possible permutations $\gamma(\mathcal{T})$ of a given feasible set $\mathcal{T}$ be denoted by $A(\mathcal{T})$. Notice that $A(\mathcal{T}) = \text{co}(\{v\}, \{e(i) \mid i \in \mathcal{T}\})$. Then the set S^n is the union of $A(\mathcal{T})$ over all feasible sets $\mathcal{T}$. The V-triangulation of S^n with grid size m^{-1} is the union of the collections $\mathcal{V}(\gamma(\mathcal{T}))$ over all possible permutations $\gamma(\mathcal{T})$ of $\mathcal{T}$ and over all feasible subsets $\mathcal{T}$. The subsets $A(\gamma(\mathcal{T}))$ and the V-triangulation are illustrated in Figure 3.2.6 for the two-dimensional case.

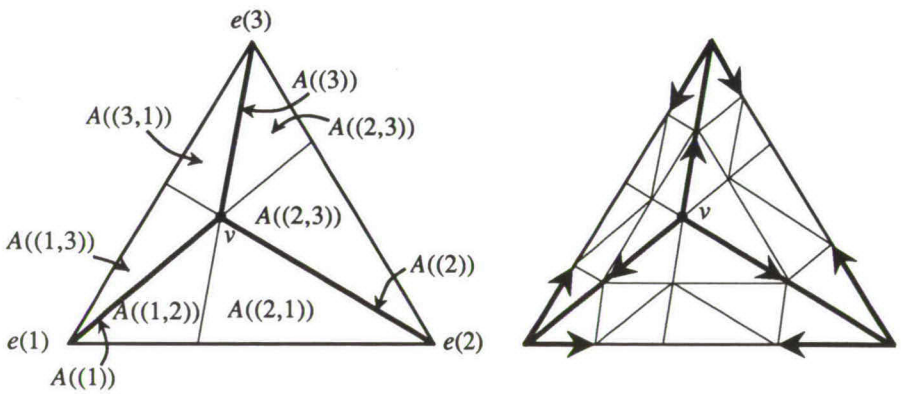


FIGURE 3.2.6: The subdivision of S^2 into subsets $A(\gamma(\mathcal{T}))$ and the V-triangulation of S^2 with grid size $\frac{1}{2}$.

The consequences of making a replacement step between two adjacent t-simplices in $\mathcal{V}(\gamma(\mathcal{T}))$ are similar to the replacement steps applicable to the Q-triangulation as $\mathcal{V}(\gamma(\mathcal{T}))$ is equivalent to the Q-triangulation of $A(\gamma(\mathcal{T}))$. Table 3.2.2 summarizes the consequences of making a replacement step in $\mathcal{V}(\gamma(\mathcal{T}))$ with the vertex y^p of σ on

the parameters y^1 and π of σ. The vector d in Table 3.2.2 is such that $d_i = d(i)$ for $i \in \mathcal{T}$ and $d_i = 0$ for $i \notin \mathcal{T}$.

TABLE 3.2.2: Replacement step in $\mathcal{V}(\gamma(\mathcal{T}))$.

	$\overline{y}^1$	$\overline{\pi}$	$\overline{d}$
$p = 1$	$y^1 + m^{-1}q(\pi_1)$	$(\pi_2, \ldots, \pi_t, \pi_1)$	$d + e(\pi_1)$
$1 < p < t+1$	y^1	$(\pi_1, \ldots, \pi_{p-2}, \pi_p, \pi_{p-1}, \pi_{p+1}, \ldots, \pi_t)$	d
$p = t+1$	$y^1 - m^{-1}q(\pi_t)$	$(\pi_t, \pi_1, \ldots, \pi_{t-1})$	$d - e(\pi_t)$

K'-triangulation of $\mathbf{R}^n$. The K'-triangulation of $\mathbf{R}^n$ was introduced in Todd (1978). To describe the K'-triangulation of $\mathbf{R}^n$ let $\mathbf{R}^n$ be subdivided into subsets $\mathbf{R}^n(s) := \text{cl}(\{x \in \mathbf{R}^n \mid \text{sgn}(x) = s\}$ where $s \in \mathbf{R}^n$ is a sign vector containing no zeros. Then $\mathbf{R}^n(s)$ is subdivided into simplices in the following way.

Definition 3.2.6 *Let s be a sign vector containing no zeros. The $K'(s)$-*triangulation *of $\mathbf{R}^n(s)$ is the collection $\mathcal{K}'(s)$ of all n-simplices $\sigma(y^1, \pi)$ with vertices $y^1, \ldots, y^{n+1}$ in $\mathbf{R}^n(s)$ such that*

i) each component of y^1 is an integer;

ii) $\pi = (\pi_1, \ldots, \pi_n)$ is a permutation of the elements of $\mathcal{I}_n$;

iii) $y^{i+1} = y^i + s_{\pi_i} e(\pi_i)$, $i = 1, \ldots, n$.

The K'-triangulation of $\mathbf{R}^n$ is the union of the collections $\mathcal{K}'(s)$ over all nonzero sign vectors s in $\mathbf{R}^n$. Figure 3.2.7 illustrates the K'-triangulation for the two-dimensional case. Notice that the K'-triangulation defined above has a grid size equal to one. For a grid size m^{-1} one needs to multiply the components of the grid points by m^{-1}. Table 3.2.1 also gives the consequences of making a replacement step on the parameters y^1 and π of $\sigma(y^1, \pi)$ in $\mathcal{K}'(s)$.

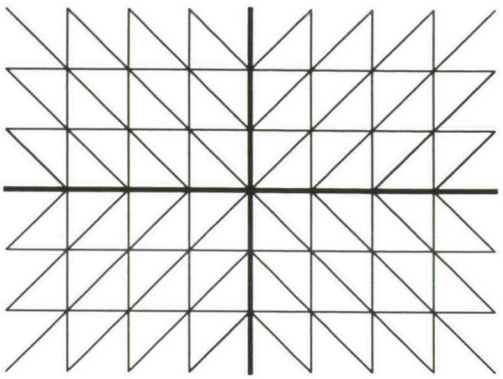


FIGURE 3.2.7: The K'-triangulation of $\mathbf{R}^2$.

D_1-triangulation of $\mathbf{R}^n$. The D_1-triangulation of $\mathbf{R}^n$ was introduced in Dang (1991a). According to measures of efficiency the D_1-triangulation is superior to the K'-triangulation of $\mathbf{R}^n$, see also Dang (1991b). It is defined as follows.

Definition 3.2.7 *The D_1-triangulation of $\mathbf{R}^n$ is the collection $\mathcal{D}$ of n-simplices $\sigma(y, \pi, s, k)$ with vertices $y^1, \ldots, y^{n+1}$ such that*

i) $1 \leq k \leq n$;

ii) $y \in \mathbf{R}^n$ *is a vector with even components;*

iii) $\pi = (\pi_1, \ldots, \pi_n)$ *is a permutation of the elements of* $\mathcal{I}_n$;

iv) $s \in \mathbf{R}^n$ *is a sign vector with components* $s_i \in \{-1, +1\}$;

v) if $k = 1$ *then* $y^1 = y$ *and* $y^{i+1} = y + s_{\pi_i} e(\pi_i)$, $i = 1, \ldots, n$,
if $k \geq 2$ *then* $y^1 = y + s$,

$$y^{i+1} = y^i - s_{\pi_i} e(\pi_i), \ i = 1, 2, \ldots, k-2, \ \textit{and}$$
$$y^{i+1} = y + s_{\pi_i} e(\pi_i), \ i = k-1, k, \ldots, n.$$

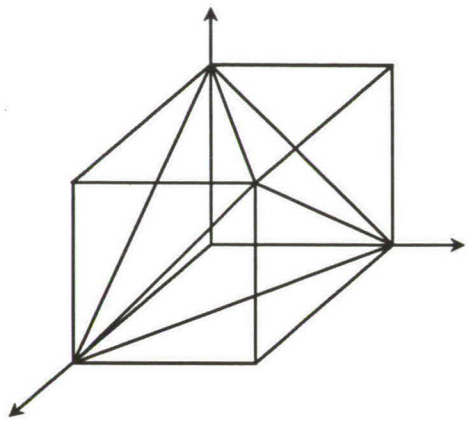


FIGURE 3.2.8: The D_1-triangulation of $\mathbf{R}^3$.

Figure 3.2.8 illustrates the D_1-triangulation for the three-dimensional case. To make a replacement step in the D_1-triangulation of $\mathbf{R}^n$ one has to apply rules different from the simplicial subdivisions above. For the exact details with respect to making a replacement step in the D_1-triangulation we refer the interested reader to Dang (1991).

3.3 Linearizations

The pivot procedure described in Section 1 can only be applied to follow a line segment in systems of linear equations like (3.1). Many of the problems considered in this monograph however contain nonlinearities. Application of the pivot procedure as described in the first section therefore often requires these nonlinearities to be replaced by a linear approximation. This section describes two possible linearizations which will be used throughout this monograph.

Let $f : \mathbf{R}^n \to \mathbf{R}^n$ be some differentiable function. The differentiability of f implies that the matrix of first-order derivatives at each point $x \in \mathbf{R}^n$ exists and is

equal to $\left[\frac{\partial f(x)}{\partial x_1}, \ldots, \frac{\partial f(x)}{\partial x_n}\right]$. This matrix is also called the *Jacobian* of f in x and is denoted by $\nabla f(x)$. Given any point $\hat{x} \in \mathbf{R}^n$ it holds for every $y \in \mathbf{R}^n \setminus \{0\}$ that $\frac{1}{\alpha} \parallel f(\hat{x} + \alpha y) - f(\hat{x}) - \alpha(\nabla f(\hat{x}))y \parallel$ tends to zero in the limit when α tends to zero. Therefore we can approximate $f(\hat{x} + \alpha y)$ for any y by $f(\hat{x}) + \alpha(\nabla f(\hat{x}))y$ for $\mid \alpha \mid$ small enough. If we replace y by $x - \hat{x}$ and take $\alpha = 1$ then the approximation $f(\hat{x}) + (\nabla f(\hat{x}))(x - \hat{x})$ is called the *first-order Taylor expansion* of f in $\hat{x}$ at a point $x \in \mathbf{R}^n$.

Definition 3.3.1 *Let f be a differentiable function from $\mathbf{R}^n$ to $\mathbf{R}^n$. Then the* first-order Taylor-expansion *of f in $\hat{x} \in \mathbf{R}^n$ at a point $x \in \mathbf{R}^n$ is given by the vector* $f(\hat{x}) + (\nabla f(\hat{x}))(x - \hat{x})$.

The first-order Taylor expansion of f in $\hat{x}$ at x is a good approximation of $f(x)$ as long as x is close enough to $\hat{x}$. Since the approximation of f may be poor in case x does not lie close to $\hat{x}$ this limits the possible use of a first-order Taylor expansion to cases where the solution of the problem under consideration does not lie in the neighbourhood of $\hat{x}$. Linearizing f by its first-order Taylor expansion in a point $\hat{x} \in \mathbf{R}^n$ is therefore a local approximation and requires f to be differentiable.

Another approximation of f is a piecewise linearization of f on a simplicial subdivision of $\mathbf{R}^n$. If one linearizes f on each n-simplex in the simplicial subdivision of $\mathbf{R}^n$ then this approximation of f, denoted F, is *piecewise linear*. Let $\mathcal{G}$ be a simplicial subdivision of $\mathbf{R}^n$. Then the *piecewise linear approximation* F of f in a point $x \in \mathbf{R}^n$ with respect to $\mathcal{G}$ is defined as follows.

Definition 3.3.2 *The* piecewise linear approximation *F of f with respect to $\mathcal{G}$ at a point $x \in \mathbf{R}^n$ is given by*

$$F(x) = \sum_{i=1}^{n+1} \lambda_i f(y^i) \tag{3.10}$$

where the convex hull $\sigma(y^1, \ldots, y^{n+1})$ of $y^1, \ldots, y^{n+1}$ in $\mathbf{R}^n$ is an n-simplex in $\mathcal{G}$ with vertices $y^1, \ldots, y^{n+1}$ containing x and where $\lambda_1, \ldots, \lambda_{n+1} \geq 0$ are such that $x = \sum_{i=1}^{n+1} \lambda_i y^i$ and $\sum_{i=1}^{n+1} \lambda_i = 1$.

In this way f is linearized on each n-simplex in $\mathcal{G}$, making F piecewise linear on $\mathbf{R}^n$.

One of the properties of a simplicial subdivision is the possibility to generate a sequence of adjacent n-simplices in $\mathcal{G}$ by making replacement steps. Let σ and $\overline{\sigma}$ be two adjacent n-simplices in $\mathcal{G}$. The linear approximation F of f on $\overline{\sigma}$ is obtained from the linear approximation F of f on σ by making a replacement step with the vertex of σ not contained in the common facet of σ and $\overline{\sigma}$.

The approximation of f by F is continually adapted when generating a sequence of adjacent n-simplices in the underlying simplicial subdivision of $\mathbf{R}^n$. This makes the quality of the approximation independent of the location in $\mathbf{R}^n$ and it is therefore not necessary for an algorithm using simplicial subdivisions to start in a neighbourhood of a solution to the problem. Hence linearizing f by its piecewise linear approximation is a global approximation. Moreover, f need not to be differentiable or even continuous. Furthermore the approximation of f by F can be improved by using a simplicial subdivision with a smaller mesh.

3.4 Homotopy interpretations

Each problem presented in this monograph can be shown to be equivalent to the stationary point problem of some function g from $\mathcal{P}$ to $\mathbf{R}^k$ where $\mathcal{P}$ is a convex and closed set in $\mathbf{R}^k$. To solve such a stationary point problem we introduce a path-following algorithm which starts in some point $x^0 \in \mathcal{P}$ and follows a path of points $x \in \mathcal{P}$ either towards a solution of the stationary point problem or diverging towards infinity[1]. The points x on the path can all be seen as a stationary point of a linear approximation G of g on some parametrized set $\mathcal{P}(\eta)$ defined as

$$\mathcal{P}(\eta) = (1 - \eta)\{x^0\} + \eta\mathcal{Q} \tag{3.11}$$

for some parameter η, $\eta \geq 0$, where $\mathcal{Q}$ is a compact and convex set in the affine hull of $\mathcal{P}$ containing x^0 in its interior. If $\eta = 0$ then $\mathcal{P}(0) = \{x^0\}$ and x^0 is a trivial solution to the stationary point problem of G on $\mathcal{P}(0)$. In case $\mathcal{P}$ is bounded then $\mathcal{P}(\eta)$ will contain $\mathcal{P}$ for η large enough and the path of stationary points of G on $\mathcal{P}(\eta)$, $\eta \geq 0$, terminates with a stationary point of G on $\mathcal{P}$. In several applications we take $\mathcal{P}$ equal to $\mathcal{Q}$ and so $\mathcal{P}(1) = \mathcal{P}$. In this case a stationary point of g on $\mathcal{P}$

[1]The algorithms presented in this monograph are shown not to cycle.

is obtained when η becomes 1. If $\mathcal{P}$ is unbounded then η may go towards infinity causing the algorithm to diverge. In that case a convergence condition on g is needed to avoid diverging.

In (3.11) we continuously deform a trivial system with a known unique solution for $\eta = 0$, the stationary point problem of g on $\{x^0\}$, into the problem we are trying to solve, the stationary point problem of g on $\mathcal{P}$. The stationary point problem of g on $\mathcal{P}(\eta)$, $0 \leq \eta$, provides a *homotopy* or *continuous deformation* of the trivial stationary point problem of g on $\{x^0\}$ into the original stationary point problem of g on $\mathcal{P}$. The parameter η is called the *homotopy parameter.*

Part B

Complementarity problems

Chapter 4

The linear complementarity problem

The linear complementarity problem is one of the most well-known problems in mathematical programming. It is frequently met in different areas of research where optimization plays an important role like operations research and economics. Section 1 of this chapter illustrates the application of the linear complementarity problem to well-known problems as linear and quadratic programming.

In operations research the linear complementarity problem can be seen as the starting point for solving more general complementarity problems such as the nonlinear complementarity problem and its generalization to the incorporation of lower and upper bounds, the generalized complementarity problem. In Chapter 6 and Chapter 7 of this monograph we consider the generalized nonlinear complementarity problem and the nonlinear complementarity problem and introduce algorithms to solve these problems.

In economics the complementarity problem mainly follows from the conditions defining an equilibrium in an economy. For example the equilibrium conditions in a pure exchange economy as well as the equilibrium conditions in an exchange economy with linear production technologies take the form of a nonlinear complementarity problem. To solve these problems algorithms have been introduced based on the approximation of the obtained nonlinear complementarity problem by a sequence of linear complementarity problems. For the computation of an equilibrium in an

exchange economy with linear production technologies we mention the algorithms in Mathiesen (1985b), Eaves (1987), as well as the algorithms introduced in Part C of this monograph.

The popularity of the linear complementarity problem in many scientific areas makes it very worthwile to introduce an algorithm to compute a solution to the problem. In the past a variety of algorithms have been introduced for this purpose among which the Lemke complementary pivoting algorithm (see Lemke (1965)) is undoubtedly one of the most renowned. In Section 2 of this chapter we describe the Lemke complementary pivoting algorithm in the framework introduced in Chapter 3. In Section 3 we introduce the algorithm as described in Talman and Van der Heyden (1983) as an extension of the Lemke complementary pivoting algorithm allowing for the possibility to choose an arbitrary starting point. In Chapter 5 of this monograph we introduce an alternative to the algorithm of Talman and Van der Heyden. But first we will define the linear complementarity problem in Section 1.

4.1 The linear complementarity problem

The *linear complementarity problem* is to find n-vectors w and z such that

$$\begin{aligned} w &= Mz + q \\ w^T z &= 0 \\ w \geq 0,\ z &\geq 0 \end{aligned} \tag{4.1}$$

for a given $(n \times n)$-matrix M and a given n-vector q. It is denoted by $LCP(q, M)$. This problem generalizes well-known problems like linear programming and quadratic programming.

The *linear programming problem* is to find a vector $x \in \mathbf{R}^n$ which solves the optimization problem

$$\begin{aligned} \max\ & c^T x \\ \text{s.t.}\ \ Ax &\leq b \\ x &\geq 0 \end{aligned} \tag{4.2}$$

where b is a given m-vector, c a given n-vector, and A a given $(m \times n)$-matrix. The Duality Theorem of Linear Programming implies that solving this maximization

problem is equivalent to solving the minimization problem

$$\begin{aligned} &\min\ y^T b \\ &\text{s.t.}\ \ c = A^T y - u \\ &\qquad y \geq 0,\ u \geq 0 \end{aligned} \tag{4.3}$$

where the m-vector y denotes the vector with dual variables to the constraints $Ax \leq b$ and the n-vector u denotes the n-vector with dual variables to the constraints in $x \geq 0$. Let v be the m-vector with slack variables to $Ax \leq b$. Then it follows from the complementarity property between the variables and the slacks in the optimization problems above that at a solution of (4.2) and (4.3) it must hold that

$$\begin{aligned} y^T v &= 0 \\ u^T x &= 0. \end{aligned} \tag{4.4}$$

To solve a linear programming problem as defined in (4.2) one can solve either (4.2) or (4.3) by using the well-known simplex method introduced in Dantzig (1951). An alternative way to find a solution to the linear programming problem as defined in (4.2) as well as a solution to (4.3) is to compute values for x, y, u, and v satisfying both the constraints in the optimization problems in (4.2) and (4.3) and the complementarity conditions in (4.4). More precisely, solving (4.2) and (4.3) is equivalent to finding $x,\ u \in \mathbf{R}^n$, and $y,\ v \in \mathbf{R}^m$, such that

$$\begin{aligned} &Ax + v = b \\ &A^T y - u = c \\ &y^T v = 0,\ u^T x = 0 \\ &y \geq 0,\ v \geq 0,\ u \geq 0,\ x \geq 0. \end{aligned}$$

This problem is a linear complementarity problem as follows from rewriting this system to

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0 & A^T \\ -A & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} -c \\ b \end{pmatrix}$$

$$\begin{pmatrix} u \\ v \end{pmatrix}^T \begin{pmatrix} x \\ y \end{pmatrix} = 0 \tag{4.5}$$

$$\begin{pmatrix} u \\ v \end{pmatrix} \geq 0,\ \begin{pmatrix} x \\ y \end{pmatrix} \geq 0.$$

A *quadratic programming problem* is to find a vector $x \in \mathbf{R}^n$ which solves the problem

$$\begin{aligned} \max \ & x^T Bx + c^T x \\ \text{s.t.} \ & Ax \leq b \\ & x \geq 0, \end{aligned} \tag{4.6}$$

where A is a given $(m \times n)$-matrix, B a given $(n \times n)$-matrix, b a given m-vector, and c a given n-vector.

To solve the quadratic programming problem one has to find a vector $x \in \mathbf{R}^n$ which fulfils the Karush-Kuhn-Tucker conditions to (4.6) given by

$$\begin{aligned} & (B + B^T)x + c = A^T y - u \\ & Ax + v = b \\ & y^T v = 0, \ u^T x = 0 \\ & y \geq 0, \ v \geq 0, \ u \geq 0, \ x \geq 0, \end{aligned}$$

where y denotes the n-vector of dual variables to the constraints in $Ax \leq b$, u denotes the n-vector of dual variables to the constraints in $x \geq 0$, and v denotes the m-vector of slack variables in $Ax \leq b$. Then the first-order conditions to the quadratic programming problem can be rewritten to

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -(B + B^T) & A^T \\ -A & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} -c \\ b \end{pmatrix}$$

$$\begin{pmatrix} u \\ v \end{pmatrix}^T \begin{pmatrix} x \\ y \end{pmatrix} = 0 \tag{4.7}$$

$$\begin{pmatrix} u \\ v \end{pmatrix} \geq 0, \quad \begin{pmatrix} x \\ y \end{pmatrix} \geq 0,$$

which is a linear complementarity problem.

These are just two examples frequently met in operations research where the linear complementarity problem plays a very important role. Except for these examples the linear complementarity problem also plays a role in areas like game theory and economics. For game theory we refer to the computation of a so-called Nash equilibrium in noncooperative games with two persons, see Lemke and Howson (1964).

In economics the linear complementarity problem is frequently met when computing an equilibrium in pure exchange economies or exchange economies with linear production technologies, see Mathiesen (1985b).

The examples given above indicate that the linear complementarity problem appears in a wide variety of problems. Therefore a variety of algorithms has been developed in the past among which the Lemke complementary pivoting algorithm (see Lemke (1965)) is undoubtedly one of the most renowned.

4.2 The Lemke algorithm

This section describes the Lemke complementary pivoting algorithm as introduced in Lemke (1965) in the framework described in Chapter 3 of this monograph. First the homotopy interpretation underlying the complementary pivoting algorithm is introduced. From this homotopy interpretation a system of equations is derived serving as a pivoting system to generate a piecewise linear path in $\mathbf{R}^n_+$. This path leads either towards a solution to the linear complementarity problem or towards infinity. Also the performance of the algorithm in the end points of the line segments making up the piecewise linear path is described. To conclude this section convergency issues concerning the Lemke algorithm are discussed as well as the impossibility of cycling.

It follows directly as a special case to Theorem 2.3.3 that the linear complementarity problem is equivalent to the stationary point problem on $\mathbf{R}^n_+$ with respect to the affine function g defined by $g(z) = -Mz - q$ on $\mathbf{R}^n_+$. This interpretation of the linear complementarity problem as a stationary point problem motivates to present the Lemke complementary pivoting algorithm as an algorithm generating a path of stationary points of g on a parametrized subset $\mathcal{H}(t)$ of $\mathbf{R}^n_+$ for a varying parameter $t \geq 0$. For $t \geq 0$, this set $\mathcal{H}(t)$ is defined as

$$\mathcal{H}(t) = \{z \in \mathbf{R}^n \mid e^\top z \leq t,\ z \geq 0\}.$$

The stationary point problem on the parametrized set $\mathcal{H}(t)$ can be interpreted as a homotopy problem with the homotopy parameter t running from zero to infinity. For $t = 0$ the algorithm is initiated in $z = 0$ which is a trivial solution to the stationary point problem of g on $\mathcal{H}(0) = \{0\}$. Then t is increased from zero thereby generating a path of stationary points of g on $\mathcal{H}(t)$ in $\mathbf{R}^n_+$. This path either ends up with a

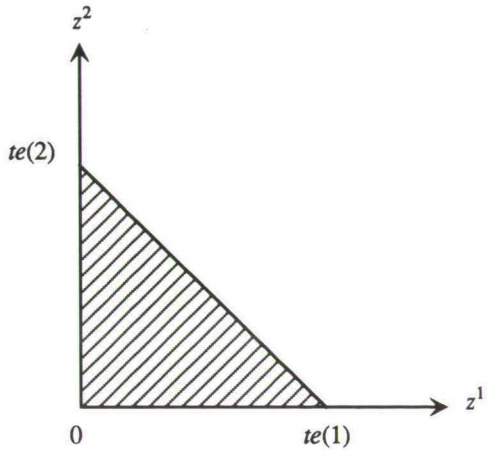


FIGURE 4.2.1: The set $\mathcal{H}(t)$ in $\mathbf{R}^2_+$ for some given positive value of the parameter t is the dashed area.

solution to the linear complementarity problem in (4.1) or diverges towards infinity. In Figure 4.2.1 we have illustrated $\mathcal{H}(t)$ for a given $t > 0$ in the two-dimensional case.

A point $\overline{z}$ in $\mathbf{R}^n_+$ is a stationary point of g on $\mathcal{H}(\overline{t})$ for some given $\overline{t} \geq 0$ if and only if $\overline{z}$ maximizes $z^\top g(\overline{z})$ over $\mathcal{H}(\overline{t})$, i.e. $\overline{z}$ solves the optimization problem, which we denote as the primal, given by

$$\begin{aligned} \max \ & z^\top g(\overline{z}) \\ \text{s.t.} \ & e^\top z \leq \overline{t} \\ & z \geq 0. \end{aligned}$$

This maximization problem is a linear programming problem. According to the Duality Theorem of Linear Programming this maximization problem is equivalent to the minimization problem, which we denote as the dual, given by

$$\begin{aligned} \min \ & \theta \overline{t} \\ \text{s.t.} \ & g(\overline{z}) = -\mu + \theta e \\ & \mu \geq 0, \ \theta \geq 0, \end{aligned}$$

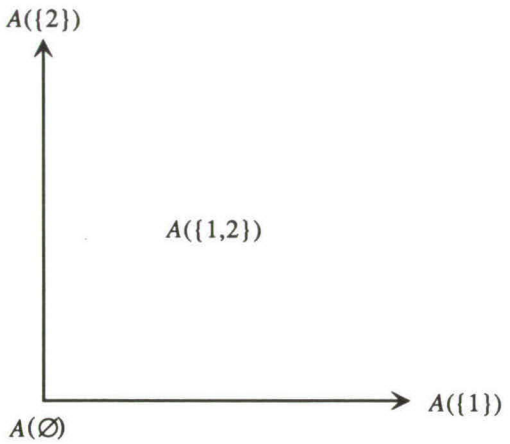


FIGURE 4.2.2: Subdivision of $\mathbf{R}^2_+$ into subsets $A(\mathcal{T})$ for $\mathcal{T} \subset \mathcal{I}_2$.

where θ is the dual variable to the constraint $e^\top z \leq \bar{t}$ of $\mathcal{H}(\bar{t})$ and μ is the n-vector containing the dual variables to the constraints in $z \geq 0$ of $\mathcal{H}(\bar{t})$ as its components.

Given $\bar{z}$ the dual has a unique solution $\bar{\theta} = \max\{\max_h g_h(\bar{z}), 0\} \geq 0$, and $\bar{\mu}_j = \bar{\theta} - g_j(\bar{z}) \geq 0$ for $j \in \mathcal{I}_n$. Let $\overline{\mathcal{T}}$ be a subset of $\mathcal{I}_n$ such that $\bar{\mu}_j = 0$ for all $j \in \overline{\mathcal{T}}$ and $\bar{\mu}_j > 0$ for all $j \notin \overline{\mathcal{T}}$. Then $\bar{\mu}_j = 0$ and $\bar{z}_j \geq 0$ if $j \in \overline{\mathcal{T}}$ and $\bar{\mu}_j > 0$ and $\bar{z}_j = 0$ if $j \notin \overline{\mathcal{T}}$. This motivates the following two definitions.

Definition 4.2.1 *For $\mathcal{T} \subseteq \mathcal{I}_n$ a point $z \in \mathbf{R}^n_+$ is $\mathcal{T}$-complete if $j \in \mathcal{T}$ implies $-(Mz)_j - q_j = \theta$ where $\theta = \max\{\max_h\{-(Mz)_h - q_h\}, 0\}$.*

Definition 4.2.2 *For $\mathcal{T} \subseteq \mathcal{I}_n$,*

$$A(\mathcal{T}) = \{z \in \mathbf{R}^n \mid z_j \geq 0 \text{ for all } j \in \mathcal{T}, \quad z_j = 0 \text{ for all } j \notin \mathcal{T}\}.$$

Figure 4.2.2 gives the subdivision of $\mathbf{R}^n_+$ into sets $A(\mathcal{T})$ for subsets $\mathcal{T}$ of $\mathcal{I}_n$ in case $n = 2$.

Theorem 4.2.1 *The point $z \in \mathbf{R}^n$ is a $\mathcal{T}$-complete point in $A(\mathcal{T})$ for some $\mathcal{T} \subseteq \mathcal{I}_n$ if and only if z is a stationary point of g on $\mathcal{H}(t)$ for some $t \geq 0$.*

Proof: Let $\overline{z}$ be a $\mathcal{T}$-complete point in $A(\mathcal{T})$ for some $\mathcal{T} \subseteq \mathcal{I}_n$. Let $\overline{t} = e^T\overline{z}$. Then $\overline{z}$ solves the primal for $t = \overline{t}$. Hence for all $z \in \mathcal{H}(\overline{t})$ it holds that $z^T g(\overline{z}) \leq \overline{z}^T g(\overline{z})$. The converse follows from the discussions above. □

In order to generate stationary points of g on $\mathcal{H}(t)$ for varying $t \geq 0$ we could generate $\mathcal{T}$-complete points in $A(\mathcal{T})$ for varying subsets $\mathcal{T}$ of $\mathcal{I}_n$.

The point z to be a $\mathcal{T}$-complete point in $A(\mathcal{T})$ for some subset $\mathcal{T}$ of $\mathcal{I}_n$ is equivalent to $\mathcal{T}$ being a subset of $\mathcal{I}_n$ such that the system of equations

$$\sum_{j \in \mathcal{T}} M_{.j} z_j - \sum_{j \notin \mathcal{T}} \mu_j e(j) + \theta e = -q \tag{4.8}$$

has a solution $z_j \geq 0$ $(j \in \mathcal{T})$, $\mu_j \geq 0$ $(j \notin \mathcal{T})$, and $\theta \geq 0$. System (4.8) is a system of n equations with $n+1$ unknowns, leaving us with one degree of freedom. If nonempty the solution set of (4.8) forms a line segment of solutions corresponding to a linear piece of $\mathcal{T}$-complete points in $A(\mathcal{T})$ with either one or two end points, for any given subset $\mathcal{T}$ of $\mathcal{I}_n$. Assuming nondegeneracy, at an end point of solutions to (4.8) exactly one of the variables is equal to zero.

The line segment of solutions to (4.8) can be followed by making a linear programming pivot step in (4.8) with one of the variables z_j $(j \in \mathcal{T})$ or μ_j $(j \notin \mathcal{T})$ or θ being zero at an end point. The other end point, if existing, corresponds to a point $\overline{z} \in \mathbf{R}^n_+$ where at the solution $(\overline{z}, \overline{\mu}, \overline{\theta})$ to (4.8) either $\overline{\theta} = 0$ or $\overline{z}_p = 0$ for some $p \in \mathcal{T}$ or $\overline{\mu}_k = 0$ for some $k \notin \mathcal{T}$. We now show that in these cases either $\overline{z}$ yields a solution to problem (4.1) or $\overline{z}$ is equal to the starting point 0 or $\overline{z}$ corresponds to an end point of a line segment of solutions to (4.8) for some other $\mathcal{T}$.

Case 1: $\overline{\theta}$ is zero. Let $\overline{w} = \sum_{j \notin \mathcal{T}} \overline{\mu}_j e(j) \geq 0$. Then $\overline{z} \geq 0$ and $\overline{w} \geq 0$ are such that $\overline{w}^T\overline{z} = 0$ and $M\overline{z} + q = \overline{w}$. Therefore $(\overline{w}, \overline{z})$ is a solution to the linear complementarity problem in (4.1).

Case 2: $\overline{z}_p$ is zero for some $p \in \mathcal{T}$. If $\mathcal{T} \setminus \{p\} = \emptyset$, then $\overline{z} = 0$. Otherwise $\overline{z}$ is a $\mathcal{T} \setminus \{p\}$-complete end point of a linear piece of $\mathcal{T} \setminus \{p\}$-complete points in $A(\mathcal{T} \setminus \{p\})$. This linear piece in $A(\mathcal{T} \setminus \{p\})$ can be generated by pivoting the column $-e(p)$ into the pivot system (4.8) thereby raising μ_p from zero and maintaining $\mathcal{T} \setminus \{p\}$-completeness.

Case 3: $\overline{\mu}_k$ is zero for some $k \notin \mathcal{T}$. Then $-(M\overline{z})_k - q_k = \overline{\theta}$ and $\overline{z}$ is a $\mathcal{T} \cup \{k\}$-complete end point of a linear piece of $\mathcal{T} \cup \{k\}$-complete points in $A(\mathcal{T} \cup \{k\})$. This

linear piece in $A(\mathcal{T} \cup \{k\})$ can be generated by pivoting the column $M_{.k}$ into (4.8) thereby raising z_k from zero and maintaining $\mathcal{T} \cup \{k\}$-completeness.

We next show that $z = 0$ is either a solution to (4.1) or an end point of a unique line segment of solutions to (4.8). If in (4.1) $\max_h -q_h \leq 0$ then $q_h \geq 0$ for all $h \in \mathcal{I}_n$. Hence $w = q \geq 0$ and $z = 0$ is a solution to the linear complementarity problem (4.1). Otherwise, let $\theta^0 = \max_h -q_h > 0$. Since we assume nondegeneracy there exists a unique $k \in \mathcal{I}_n$ such that $-q_k = \theta^0$. The point $z = 0$ is therefore $\{k\}$-complete and the system of equations

$$-\sum_{j \neq k} \mu_j e(j) + \theta e = -q \tag{4.9}$$

has a unique solution $\theta^0 = \max_h -q_h > 0$ and $\mu_j^0 = \theta^0 + q_j > 0$ $(j \neq k)$. This implies that $z = 0$ is a $\{k\}$-complete end point of a linear piece of $\{k\}$-complete points in $A(\{k\})$. This linear piece corresponds to a line segment of solutions to the system of equations

$$M_{.k} z_k - \sum_{j \neq k} \mu_j e(j) + \theta e = -q.$$

This line segment of solutions can be generated by pivoting the column $M_{.k}$ into (4.9) thereby raising z_k from zero.

The above cases show that each end point of a line segment of solutions to (4.8) either corresponds to the starting point $z = 0$ or to a solution of the linear complementarity problem (4.1) or is an end point of a line segment of solutions to exactly one other system of equations for a different subset of $\mathcal{I}_n$. The point $z = 0$ corresponds to an end point of exactly one line segment of solutions to (4.8).

To find a solution of problem (4.1) Lemke's algorithm starts in $z = 0$ and generates a sequence of adjacent line segments of solutions to (4.8) for varying $\mathcal{T} \subseteq \mathcal{I}_n$. Each line segment of solutions is generated by making a pivot step in (4.8) with one of the variables being zero at an end point until another variable becomes zero. Such a line segment of solutions to (4.8) corresponds to a linear piece of $\mathcal{T}$-complete points in $A(\mathcal{T})$.

These properties make the path of points generated by the algorithm a piecewise linear path through subsequent subsets $A(\mathcal{T})$ of $\mathbf{R}_+^n$ for varying $\mathcal{T} \subseteq \mathcal{I}_n$. As we show

next, this path either ends in a solution to the linear complementarity problem or in a ray of solutions to (4.8), but it cannot cycle.

Theorem 4.2.2 *The Lemke complementary pivoting algorithm cannot cycle.*

Proof: Suppose Lemke's algorithm cycles and let $\overline{z}$ be the first end point generated twice. If $\overline{z} \neq 0$, then $\overline{z}$ corresponds to an end point of a line segment of solutions to (4.8) for some subset $\mathcal{T}$ of $\mathcal{I}_n$ and to an end point of a line segment of solutions to (4.8) for some subset $\overline{\mathcal{T}}$ of $\mathcal{I}_n$, $\mathcal{T} \neq \overline{\mathcal{T}}$. With respect to the set $\mathcal{T}$, in $\overline{z}$ either Case 2 or Case 3 occurs by the assumption of nondegeneracy. If Case 2 occurs then $\overline{z}$ is also an end point of a line segment of solutions to (4.8) for $\mathcal{T} \setminus \{p\}$. The assumption of nondegeneracy implies that $\overline{\mathcal{T}} = \mathcal{T} \setminus \{p\}$. Hence the algorithm returns to $\overline{z}$ after having made the next pivoting step in (4.8). Otherwise $\overline{z}$ was not the first end point to be visited twice. But this contradicts the fact that the algorithm never goes back at an end point.

In a similar way it can be proved that an end point $\overline{z} \neq 0$ cannot be generated for a second time if Case 3 occurs. Therefore, $\overline{z} = 0$. As $\overline{z} = 0$ is an end point of a line segment of solutions to (4.8) for a unique $\mathcal{T}$ and the algorithm starts to follow this line segment, the algorithm cannot return to $\overline{z}$ either. □

Since the number of subsets $\mathcal{T}$ of $\mathcal{I}_n$ is finite Theorem 4.2.2 causes the algorithm either to end up with a solution to the linear complementarity problem in (4.1) or in a ray of solutions to (4.8) for some subset $\mathcal{T}$ of $\mathcal{I}_n$. The convergency results on the Lemke complementary pivoting algorithm are therefore concerned with conditions on M and q under which no ray of solutions to (4.8) for some subset $\mathcal{T}$ of $\mathcal{I}_n$ is generated. The weakest conditions on the convergence of the Lemke algorithm until now have been given in Jones (1986). We summarize his results in the next theorem. We remark that a square matrix C is said to be *copositive* if $x^\top Cx \geq 0$ whenever x is nonnegative, and a square matrix P is said to be *copositive-plus* if P is copositive and if, in addition, $(P + P^\top)z = 0$ whenever $z^\top Pz = 0$, $z \geq 0$.

Theorem 4.2.3 *Suppose M can be written as $P + C$ where P is copositive-plus and symmetric and C is copositive. If the system $q + Px - C^\top y \geq 0$, $y \geq 0$, is feasible, then the algorithm terminates in a solution.*

Proof: See Jones (1986). □

Notice that in Lemke (1965) the algorithm is described as an algorithm following a piecewise linear path of complementary points in $\mathbf{R}^{2n}$. Our description given above coincides with his as can easily be seen.

In a point $\overline{z}$ on the path there exists a subset $\mathcal{T}$ of $\mathcal{I}_n$ such that $\overline{z}$ is a $\mathcal{T}$-complete point in $A(\mathcal{T})$. By definition of $\mathcal{T}$-completeness this means that $(M\overline{z})_j + q_j + \theta \geq 0$ and $\overline{z}_j = 0$ or $(M\overline{z})_j + q_j + \theta = 0$ and $\overline{z}_j \geq 0$ for every $j \in \mathcal{I}_n$. Hence $\overline{z}$ and $M\overline{z} + q + \theta e$ are complementary. In an end point of a line segment of $\mathcal{T}$-complete points on the path there exists an $h \in \mathcal{I}_n$ such that $(M\overline{z})_h + q_h + \theta = 0$ and $\overline{z}_h = 0$ while for all $j \in \mathcal{I}_n \setminus \{h\}$ either $(M\overline{z})_j + q_j + \theta > 0$ and $\overline{z}_j = 0$ or $(M\overline{z})_j + q_j + \theta = 0$ and $\overline{z}_j > 0$. Lemke refers to such points as *almost-complementary* points. The procedure of the algorithm in these end points as described in the cases above is therefore called *complementary pivoting.* When μ_k becomes zero in an end point the next pivot step is made in (4.8) with its complementary variable z_k and when z_k becomes zero the next pivot step is made with μ_k.

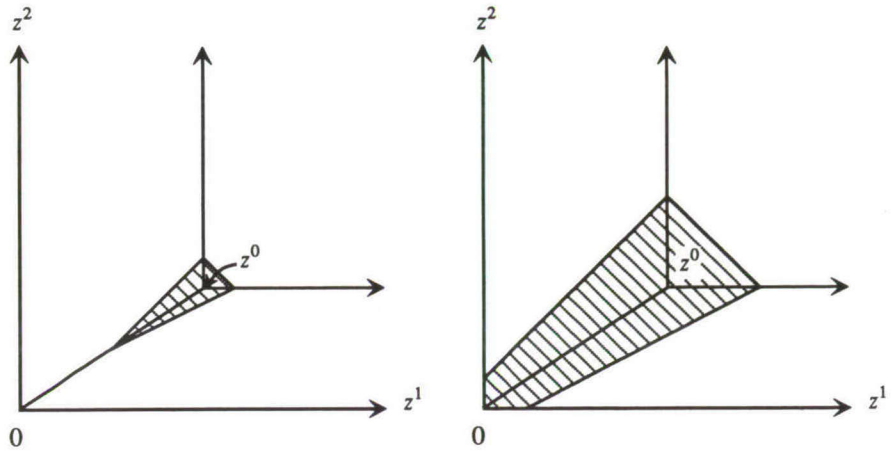
4.3 The Talman and Van der Heyden algorithm

In Talman and Van der Heyden (1983) the Lemke complementary pivoting algorithm presented in the previous section of this chapter is generalized by incorporating the possibility of an arbitrary starting point in the algorithm. For this purpose they introduce a whole class of algorithms all of which incorporate the starting point in a different way. In this section we only give a review of the most efficient one in a slightly adjusted way.

The algorithm as proposed in Talman and Van der Heyden (1983) follows a path of points in $\mathbf{R}^n_+$ starting in some arbitrarily chosen point $z^0 \in \text{int}(\mathbf{R}^n_+)$. Each point z on the path is such that it is a stationary point of the affine function g defined in the previous section on the set $\mathcal{H}(t)$, for some $t \geq 0$, where

$$\mathcal{H}(t) := \{(1 - \lambda_{n+1})z^0 + \sum_{j=1}^{n} \lambda_j e(j) \mid \lambda_j \geq 0 \ (j \in \mathcal{I}_n),\ 0 \leq \lambda_{n+1} \leq 1, \text{ and } \sum_{j=1}^{n+1} \lambda_j \leq t\}.$$

The algorithm is initiated in z^0 which is a trivial solution to the stationary point



(a) $\mathcal{H}(\frac{1}{2})$ (b) $\mathcal{H}(\frac{3}{2})$

FIGURE 4.3.1: The set $\mathcal{H}(t)$ in $\mathbf{R}^2_+$ for $t = \frac{1}{2}$ and $t = \frac{3}{2}$, given $z^0 = (3,2)^\top$.

problem of g on $\mathcal{H}(0) = \{z^0\}$. Then the algorithm raises t from zero thereby generating a piecewise linear path of points in $\mathbf{R}^n_+$ which either ends up with a solution to the linear complementarity problem (4.1) or in a ray in which case no solution to the linear complementarity problem will be found. Figure 4.3.1 illustrates the set $\mathcal{H}(t)$ for two positive values of t in the two-dimensional case.

A point $\overline{z}$ in $\mathbf{R}^n_+$ is a stationary point of g on $\mathcal{H}(\overline{t})$ for some $\overline{t} \geq 0$ if and only if $\overline{z}$ maximizes $z^\top g(\overline{z})$ over $z \in \mathcal{H}(\overline{t})$. Then by definition of $\mathcal{H}(t)$ there exist numbers $\overline{\lambda}_1, \ldots, \overline{\lambda}_{n+1}$ such that $\overline{z} = (1 - \overline{\lambda}_{n+1})z^0 + \sum_{j=1}^n \overline{\lambda}_j e(j)$ which solve the maximization problem, called the primal, given by

$$\begin{aligned} \max \ & \textstyle\sum_{i=1}^n \lambda_i g_i(\overline{z}) - \lambda_{n+1} z^{0\top} g(\overline{z}) \\ \text{s.t.} \ & \lambda_i \geq 0 \ (i \in \mathcal{I}_n) \\ & 0 \leq \lambda_{n+1} \leq 1 \\ & \textstyle\sum_{i=1}^{n+1} \lambda_i \leq \overline{t}. \end{aligned}$$

This maximization problem is a linear programming problem. According to the Duality Theorem of Linear Programming, this maximization problem is equivalent

to the minimization problem, called the dual, given by

$$\begin{aligned} \min\ & \theta\bar{t} + \xi_2 \\ \text{s.t.}\ & g_i(\bar{z}) = -\mu_i + \theta\ (i \in \mathcal{I}_n) \\ & -z^{0\mathsf{T}} g(\bar{z}) = \xi_2 - \xi_1 + \theta \\ & \theta \geq 0,\ \mu_i \geq 0\ (i \in \mathcal{I}_n), \end{aligned}$$

where θ is the dual variable to the constraint $\sum_{i=1}^{n+1} \lambda_i \leq \bar{t}$, μ_i $(i \in \mathcal{I}_n)$ are the dual variables to the constraints $\lambda_i \geq 0$ $(i \in \mathcal{I}_n)$, ξ_1 is the dual variable to the constraint $-\lambda_{n+1} \geq 0$, and ξ_2 is the dual variable to the constraint $\lambda_{n+1} \leq 1$.

Given $\bar{z} = (1 - \bar{\lambda}_{n+1})z^0 + \sum_{j=1}^{n} \bar{\lambda}_j e(j)$ the dual has a unique solution $\bar{\theta} = \max\{\max_h g_h(\bar{z}),\ -z^{0\mathsf{T}} g(\bar{z})\}$ and $\bar{\xi}_2 = 0$ if $\bar{\lambda}_{n+1} < 1$, $\bar{\theta} = \max\{\max_h g_h(\bar{z}),\ 0\}$ and $\bar{\xi}_2 = -z^{0\mathsf{T}} g(\bar{z}) - \bar{\theta}$ if $\bar{\lambda}_{n+1} = 1$, $\bar{\mu}_j = \bar{\theta} - g_j(\bar{z}) \geq 0$, and $\bar{\xi}_1 = \bar{\xi}_2 + \bar{\theta} + z^{0\mathsf{T}} g(\bar{z})$. Let $\overline{\mathcal{T}}$ be a subset of $\mathcal{I}_{n+1}$ such that $\bar{\mu}_j = 0$ for all $j \in \overline{\mathcal{T}}$, $\bar{\mu}_j > 0$ for all $j \notin \overline{\mathcal{T}} \cup \{n+1\}$, and $\bar{\xi}_2 - \bar{\xi}_1 \neq 0$ if $n+1 \notin \overline{\mathcal{T}}$. Then $\bar{\mu}_j = 0$ and $\bar{\lambda}_j \geq 0$ if $j \in \overline{\mathcal{T}} \setminus \{n+1\}$, $\bar{\mu}_j > 0$ and $\bar{\lambda}_j = 0$ if $j \notin \overline{\mathcal{T}} \cup \{n+1\}$, $\bar{\xi}_1 = \bar{\xi}_2 = 0$ and $0 < \bar{\lambda}_{n+1} < 1$ if $n+1 \in \overline{\mathcal{T}}$ and either $\bar{\xi}_1 = 0$, $\bar{\xi}_2 > 0$, and $\bar{\lambda}_{n+1} = 1$, or $\bar{\xi}_1 > 0$, $\bar{\xi}_2 = 0$, and $\bar{\lambda}_{n+1} = 0$ if $n+1 \notin \overline{\mathcal{T}}$. Define the variable μ_{n+1} to be equal to $\xi_2 - \xi_1$. This leads to the following definitions.

Definition 4.3.1 *For $\mathcal{T} \subset \mathcal{I}_{n+1}$ a point $z \in \mathbf{R}^n_+$ is $\mathcal{T}$-complete if $j \in \mathcal{T}$ implies $-(Mz)_j - q_j = \theta$ and $n+1 \in \mathcal{T}$ when $z^{0\mathsf{T}}(Mz+q) = \theta$ where if $z \in \mathrm{bd}(\mathbf{R}^n_+)$, then $\theta = \max\{\max\{-(Mz)_h - q_h\},\ 0\}$ and if $z \in \mathbf{R}^n_{++}$*

$$\theta = \max \begin{cases} -(Mz)_h - q_h, & h \in \mathcal{I}_n \\ z^{0\mathsf{T}}(Mz+q). \end{cases}$$

Definition 4.3.2 *For $\mathcal{T} \subset \mathcal{I}_{n+1}$, $\mathcal{T} \neq \emptyset$*

$$A(\mathcal{T}) = \{(1-\lambda_{n+1})z^0 + \sum_{j \in \mathcal{T} \setminus \{n+1\}} \lambda_j e(j) \mid \lambda_j \geq 0,\ j \in \mathcal{T} \setminus \{n+1\},\ 0 \leq \lambda_{n+1} \leq 1\}.$$

Definition 4.3.3 *For $\mathcal{T} \subset \mathcal{I}_n$ such that $\mathcal{T} \neq \mathcal{I}_n$*

$$A^0(\mathcal{T}) = \{\sum_{j \in \mathcal{T}} \lambda_j e(j) \mid \lambda_j \geq 0,\ j \in \mathcal{T}\}.$$

Figure 4.3.2 illustrates the subdivision of $\mathbf{R}^n_+$ into subsets $A(\mathcal{T})$ and $A^0(\mathcal{T})$ when $n = 2$ and $z^0 = (3,2)^{\mathsf{T}}$. Notice that $A^0(\emptyset) = \{0\}$.

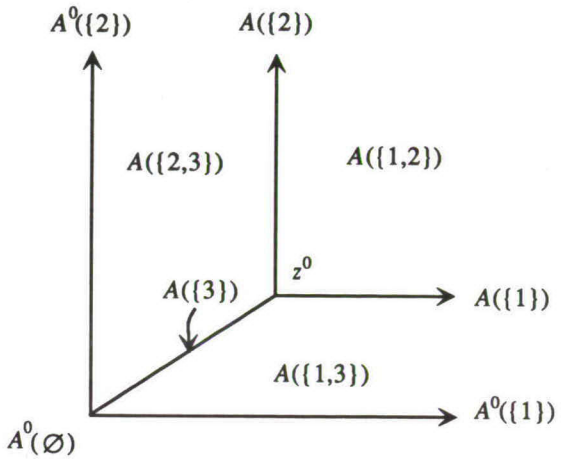


FIGURE 4.3.2: Subdivision of $\mathbf{R}^2_+$ into subsets $A(\mathcal{T})$ and $A^0(\mathcal{T})$ for $\mathcal{T} \subset \mathcal{I}_3$ given the starting point $z^0 = (3,2)^\top$.

Theorem 4.3.1 *The point z is a $\mathcal{T}$-complete point in $A(\mathcal{T})$ or $A^0(\mathcal{T})$ for some feasible subset $\mathcal{T}$ of $\mathcal{I}_{n+1}$ if and only if z is a stationary point of g on $\mathcal{H}(t)$ for some $t \geq 0$.*

Proof: Similar to the proof of Theorem 4.2.1. □

Theorem 4.3.1 implies that to generate a path of stationary points of g on $\mathcal{H}(t)$ for a varying parameter $t \geq 0$ the algorithm could generate $\mathcal{T}$-complete points in $A(\mathcal{T})$ or in $A^0(\mathcal{T})$ for varying subsets $\mathcal{T}$ of $\mathcal{I}_{n+1}$.

In general the $\mathcal{T}$-completeness condition at a point z on the path is equivalent to $\mathcal{T}$ being a subset of $\mathcal{I}_{n+1}$ such that the system of equations

$$\begin{pmatrix} -Mz - q \\ z^{0\top}(Mz + q) \end{pmatrix} = \theta e - \sum_{j \notin \mathcal{T}} \mu_j e(j)$$

has a solution z, $\theta \geq 0$, $\mu_j \geq 0$ $(j \notin \mathcal{T} \cup \{n+1\})$, $\mu_{n+1} \geq 0$ if $n+1 \notin \mathcal{T}$ and $z \in A(\mathcal{T})$, $\mu_{n+1} \leq 0$ if $z \in A^0(\mathcal{T})$. Then Lemma 4.3.1 gives the appropriate pivot system to generate a linear piece of $\mathcal{T}$-complete points in $A(\mathcal{T})$, and Lemma 4.3.2

gives the appropriate pivot system to generate a linear piece of $\mathcal{T}$-complete points in $A^0(\mathcal{T})$.

Lemma 4.3.1 *For some feasible $\mathcal{T}$, a point $z \in \mathbf{R}^n$ is a $\mathcal{T}$-complete point in $A(\mathcal{T})$ if and only if the system of equations*

$$\sum_{j \in \mathcal{T} \setminus \{n+1\}} \lambda_j \begin{pmatrix} M_{.j} \\ -z^{0\mathsf{T}} M_{.j} \end{pmatrix} - \lambda_{n+1} \begin{pmatrix} Mz^0 \\ -z^{0\mathsf{T}} Mz^0 \end{pmatrix} - \sum_{j \notin \mathcal{T}} \mu_j e(j) + \theta e = \begin{pmatrix} -q - Mz^0 \\ z^{0\mathsf{T}}(q + Mz^0) \end{pmatrix} \quad (4.10)$$

has a solution $\lambda_j \geq 0$ $(j \in \mathcal{T} \setminus \{n+1\})$, $0 \leq \lambda_{n+1} \leq 1$, $\mu_j \geq 0$ $(j \notin \mathcal{T})$, $\theta \geq 0$ such that $z = (1 - \lambda_{n+1})z^0 + \sum_{j \in \mathcal{T} \setminus \{n+1\}} \lambda_j e(j)$.

Lemma 4.3.2 *For some feasible $\mathcal{T}$, a point $z \in \mathbf{R}^n$ is a $\mathcal{T}$-complete point in $A^0(\mathcal{T})$ if and only if the system of equations*

$$\sum_{j \in \mathcal{T}} \lambda_j \begin{pmatrix} M_{.j} \\ -z^{0\mathsf{T}} M_{.j} \end{pmatrix} - \sum_{j \notin \mathcal{T}} \mu_j e(j) + \theta e = \begin{pmatrix} -q \\ z^{0\mathsf{T}} q \end{pmatrix} \quad (4.11)$$

has a solution $\lambda_j \geq 0$ $(j \in \mathcal{T})$, $\mu_j \geq 0$ $(j \notin \mathcal{T} \cup \{n+1\})$, $\mu_{n+1} \leq 0$, $\theta \geq 0$ such that $z = \sum_{j \in \mathcal{T}} \lambda_j e(j)$.

The systems of equations in Lemma 4.3.1 and Lemma 4.3.2 both contain $n+1$ equations and $n+2$ variables leaving us with one degree of freedom. If nonempty, the solution set of each system forms a line segment, assuming nondegeneracy. This line segment corresponds to a linear piece of $\mathcal{T}$-complete points in $A(\mathcal{T})$ or $A^0(\mathcal{T})$ with either one or two end points.

The algorithm follows a linear piece of $\mathcal{T}$-complete points in $A(\mathcal{T})$ or $A^0(\mathcal{T})$ by making a linear programming pivot step in the appropriate system of equations with one of the variables λ_j $(j \in \mathcal{T})$ or μ_j $(j \notin \mathcal{T})$ being zero in an end point of a line segment on the path. The other end point of this line segment, if existing, is a point $\overline{z} \in \mathbf{R}^n_+$ where at the solution $(\overline{\theta}, \overline{\lambda}, \overline{\mu})$, to (4.10) or (4.11), either $\overline{\theta} = 0$, or $\overline{\lambda}_p = 0$ for some $p \in \mathcal{T}$ such that $\mathcal{T} \setminus \{p\} \neq \emptyset$, or $\overline{\lambda}_{n+1} = 1$, or $\overline{\mu}_k = 0$ for some $k \notin \mathcal{T}$.

Case 1: $\overline{\theta}$ is 0. If $\overline{z} \in A^0(\mathcal{T})$ then $\overline{z} = \sum_{j \in \mathcal{T}} \overline{\lambda}_j e(j)$ and $\overline{w} = M\overline{z} + q = \sum_{j \notin \mathcal{T}} \mu_j e(j)$ is a solution to the linear complementarity problem since $\overline{w}^{\mathsf{T}} \overline{z} = 0$. If $\overline{z} \in A(\mathcal{T})$ then $\overline{z} = (1 - \overline{\lambda}_{n+1})z^0 + \sum_{j \in \mathcal{T} \setminus \{n+1\}} \overline{\lambda}_j e(j)$ and $\overline{w} = M\overline{z} + q = \sum_{j \notin \mathcal{T}} \overline{\mu}_j e(j)$. The $(n+1)$-th equation implies that $-z^{0\mathsf{T}} \overline{w} = \overline{\mu}_{n+1} \geq 0$. Also $z^{0\mathsf{T}} \overline{w} \geq 0$ because $z^0 > 0$

and $\overline{w} \geq 0$. Therefore $z^{0\mathsf{T}}\overline{w} = 0$, and so $\overline{w} = 0$. Hence $\overline{z}$ and $\overline{w}$ is a solution to the linear complementarity problem.

Case 2: $\overline{\lambda}_p$ is 0 for some $p \in \mathcal{T}$ such that $\mathcal{T} \setminus \{p\} \neq \emptyset$. Then $\overline{z}$ is a $\mathcal{T} \setminus \{p\}$-complete end point of a linear piece of $\mathcal{T} \setminus \{p\}$-complete points in $A(\mathcal{T} \setminus \{p\})$ or in $A^0(\mathcal{T} \setminus \{p\})$ depending on whether $\overline{z} \in A(\mathcal{T})$ or $\overline{z} \in A^0(\mathcal{T})$ respectively. The algorithm proceeds in $A(\mathcal{T}\setminus\{p\})$ or $A^0(\mathcal{T}\setminus\{p\})$ by pivoting the column $-e(p)$ into the appropriate system of equations thereby raising μ_p from zero and maintaining $\mathcal{T} \setminus \{p\}$-completeness.

Case 3: $\overline{\lambda}_{n+1}$ is 1 in system (4.10). Then $\overline{z} = \sum_{j\in\mathcal{T}\setminus\{n+1\}} \overline{\lambda}_j e(j)$ in $A(\mathcal{T})$ is a $\mathcal{T} \setminus \{n+1\}$-complete end point of a linear piece of $\mathcal{T} \setminus \{n+1\}$-complete points in $A^0(\mathcal{T} \setminus \{n+1\})$. The algorithm proceeds by first moving the column corresponding to the variable $\overline{\lambda}_{n+1}$ to the right hand side and then pivoting the column $e(n+1)$ into (4.11) thereby lowering μ_{n+1} from zero. In this way the algorithm maintains $\mathcal{T} \setminus \{n+1\}$-completeness in $A^0(\mathcal{T} \setminus \{n+1\})$.

Case 4: $\overline{\mu}_k$ is 0 for some $k \notin \mathcal{T}$. If $\mathcal{T} \cup \{k\} = \mathcal{I}_{n+1}$ then $\overline{w} = M\overline{z}+q = -\overline{\theta}e \geq 0$ and $\overline{\theta} \geq 0$ implies $\overline{\theta} = 0$. Hence $\overline{w} = 0$ while $\overline{z} > 0$, and the algorithm stops with a solution to the linear complementarity problem. Otherwise $\mathcal{T} \cup \{k\} \neq \mathcal{I}_{n+1}$. If $\overline{z} \in A(\mathcal{T})$ then $\overline{z}$ is also a $\mathcal{T} \cup \{k\}$-complete end point of a linear piece of $\mathcal{T} \cup \{k\}$-complete points in $A(\mathcal{T} \cup \{k\})$. The algorithm proceeds by pivoting the column $((M_{.k})^{\mathsf{T}}, -z^{0\mathsf{T}}M_{.k})^{\mathsf{T}}$ into (4.10) thereby raising λ_k from zero and generating a linear piece of $\mathcal{T} \cup \{k\}$-complete points in $A(\mathcal{T} \cup \{k\})$. If $\overline{z} \in A^0(\mathcal{T})$ and $k \neq n+1$ then $\overline{z}$ is also a $\mathcal{T} \cup \{k\}$-complete end point of a linear piece of $\mathcal{T} \cup \{k\}$-complete points in $A^0(\mathcal{T} \cup \{k\})$. The algorithm proceeds by pivoting the column $((M_{.k})^{\mathsf{T}}, -z^{0\mathsf{T}}M_{.k})^{\mathsf{T}}$ into (4.11) thereby raising λ_k from zero. If $\overline{z} \in A^0(\mathcal{T})$ and $k = n+1$ then $\overline{z}$ is a $\mathcal{T} \cup \{n+1\}$-complete end point of a linear piece of $\mathcal{T} \cup \{n+1\}$-complete points in $A(\mathcal{T} \cup \{n+1\})$. The algorithm proceeds by first subtracting the column $((Mz^0)^{\mathsf{T}}, -z^{0\mathsf{T}}Mz^0)^{\mathsf{T}}$ from the right hand side and then pivoting it into (4.11) thereby lowering the variable λ_{n+1} from one. Hence, the algorithm generates a line segment of solutions to (4.10) thereby maintaining $\mathcal{T} \cup \{n+1\}$-completeness.

To show that z^0 is an end point of a line segment of solutions to (4.10) for some unique subset $\mathcal{T}^0$ of $\mathcal{I}_{n+1}$ let $\theta^0 = \max\{\max_h\{-(Mz^0)_h - q_h\},\ z^{0\mathsf{T}}(Mz^0+q)\}$. Since we have assumed nondegeneracy there exists either a unique k such that $\theta^0 =$

$-(Mz^0)_k - q_k$ and so $\mathcal{T}^0 = \{k\}$, or $\theta^0 = z^{0\mathsf{T}}(Mz^0 + q)$ and so $\mathcal{T}^0 = \{n+1\}$. The starting point z^0 is $\mathcal{T}^0$-complete and the system of equations

$$-\sum_{j \notin \mathcal{T}^0} \mu_j e(j) + \theta e = \begin{pmatrix} -q - Mz^0 \\ z^{0\mathsf{T}}(q + Mz^0) \end{pmatrix} \tag{4.12}$$

has a unique solution $\theta = \theta^0 > 0$ and $\mu_j = q_j + (Mz^0)_j + \theta^0 > 0$ for all $j \notin \mathcal{T}^0 \cup \{n+1\}$ and $\mu_{n+1} = \theta^0 - z^{0\mathsf{T}}(q + Mz^0) > 0$ if $\{n+1\} \notin \mathcal{T}^0$. This implies that z^0 is an end point of a linear piece of $\mathcal{T}^0$-complete points in $A(\mathcal{T}^0)$. In order to follow this linear piece of $\mathcal{T}^0$-complete points in $A(\mathcal{T}^0)$ the algorithm starts by pivoting the column $((M_{.k})^\mathsf{T},\ -(M_{.k})^\mathsf{T} z^0)^\mathsf{T}$ into (4.12) if $\mathcal{T}^0 = \{k\}$ with $k < n+1$ thereby raising λ_k from zero and the column $(-z^{0\mathsf{T}} M^\mathsf{T},\ z^{0\mathsf{T}} M z^0)^\mathsf{T}$ if $\mathcal{T}^0 = \{n+1\}$ into (4.12) thereby raising λ_{n+1} from zero.

The above cases indicate that each end point of a line segment of solutions to (4.10) or (4.11) either corresponds to the starting point z^0 or to a solution of the linear complementarity problem or is an end point of a line segment of solutions to exactly one other system of equations for a different subset $\mathcal{T}$ of $\mathcal{I}_{n+1}$. The point z^0 corresponds to an end point of exactly one line segment of solutions to (4.10).

These properties make the path of points generated by the algorithm a piecewise linear path of points through subsequent subsets $A(\mathcal{T})$ and $A^0(\mathcal{T})$ of $\mathbf{R}^n_+$ for varying subsets $\mathcal{T}$ of $\mathcal{I}_{n+1}$. Each linear piece of this path is followed by making a linear programming pivot step in the appropriate system of equations with a variable being zero or making a binding constraint at an end point. In this way the algorithm may either end up with a solution to the linear complementarity problem, or in a ray of solutions to the appropriate system of equations. Similar to Theorem 4.2.2 for the Lemke complementary pivoting algorithm one can prove that this algorithm does not cycle either.

During the enumeration of the cases above two possibilities of the algorithm stopping in a solution occurred. The next lemma summarizes these cases.

Lemma 4.3.3 *Let z be an end point of a line segment of $\mathcal{T}$-complete points in $A(\mathcal{T})$ or $A^0(\mathcal{T})$ for some subset $\mathcal{T}$ of $\mathcal{I}_{n+1}$. Then z is a solution to the linear complementarity problem if*

i) $\theta = 0$;

ii) $\mu_k = 0$, *for some* $k \notin \mathcal{T}$, *and* $\mathcal{T} \cup \{k\} = \mathcal{I}_{n+1}$.

Notice that case ii) in Lemma 4.3.3 implies that $\theta = 0$. Therefore at the end point z in which this case is valid the system of equations (4.11) is degenerate.

If the algorithm does not find a solution it ends up in a ray of solutions to one of the systems of equations in Lemma 4.3.1 and Lemma 4.3.2. To find conditions under which the algorithm converges is therefore concerned with conditions under which the algorithm does not generate such a ray of solutions. It can be shown that the algorithm converges under the same conditions under which the Lemke complementary pivoting algorithm converges as given in Theorem 4.2.2, see Jones (1986).

Chapter 5

A new pivoting algorithm to solve LCP's

The popularity of the linear complementarity problem in mathematical programming has led to a variety of algorithms attempting to solve the problem. Among this variety of algorithms the Lemke complementary pivot algorithm presented in Chapter 4 is undoubtedly one of the most renowned algorithms. As we saw in Chapter 4 the Lemke complementary pivoting algorithm is a path-following algorithm starting in $z = 0$ and generating a piecewise linear path either towards a solution to the linear complementarity problem or towards infinity.

The major drawback of the Lemke complementary pivoting algorithm however is that one is stuck to the fixed starting point $z = 0$, a feature which causes inefficiencies when one has some idea concerning the possible location of a solution to the linear complementarity problem. Such information might for example be available in case one tries to solve a nonlinear complementarity problem by a sequence of linear complementarity problems. This inefficiency in processing the information makes it worthwile to adapt Lemke's algorithm for an arbitrarily chosen starting point. Talman and Van der Heyden (1983) present a whole class of algorithms extending the Lemke complementary pivoting algorithm to the possibility to start in an arbitrarily chosen starting point. We have introduced a slightly adapted version of the most efficient one in this class in Section 3 of the previous chapter. All the algorithms in this class however use a pivot system of at least $n+1$ equations in order to guarantee

possible convergence of the algorithm where n is the dimension of the linear complementarity problem. Moreover none of these algorithms seem to be very natural in solving the linear complementarity problem. In this chapter we propose a new pivoting algorithm to solve the linear complementarity problem allowing for an arbitrarily chosen starting point. This algorithm has a natural interpretation as a path-following algorithm and it does not need more than n equations in the pivot system.

The algorithm leaves the starting point in one out of $n+1$ possible directions. There are n rays that connect the starting point with each of the n axes of $\mathbf{R}^n_+$ and one ray that connects the starting point with the origin. This allows the algorithm to leave the starting point z^0 in such a way that, with $w^0 = Mz^0 + q$, it will raise z_i from z_i^0 when w_i^0 is negative and smaller than all other components of w^0, while the algorithm will lower each z_i from z_i^0 proportionally towards zero when w_i^0 is positive or not smaller than all other components of w^0. In particular this latter feature endows the algorithm with a very natural interpretation. For example, the algorithm will stop with a solution to the linear complementarity problem if it reaches the origin. This is contrary to the algorithm in the Talman and Van der Heyden class of algorithms presented in the previous chapter having also $n+1$ rays to leave the starting point. In that algorithm there are n rays that leave the starting point parallel to each of the axes and there is one ray connecting the starting point with the origin. The algorithm may be forced to continue along one of the axes of $\mathbf{R}^n_+$ when reaching the origin.

In our algorithm the intersection of the rays with each of the axes can arbitrarily be chosen. In Section 3 of this chapter we suggest a particular choice of these intersections such that it is possible to see in advance whether the algorithm might not solve the problem.

This chapter is based on Kremers and Talman (1990c) and is divided as follows. First we describe the algorithm. The algorithm follows a path of points in $\mathbf{R}^n_+$. Each point on this path can be interpreted as a stationary point of the affine function g already defined in Chapter 4 on a parametrized set. This interpretation is elaborated in Section 1. The steps of the algorithm are enumerated in Section 2 while Section 3 is dedicated to convergency issues.

5.1 The algorithm

The algorithm follows a piecewise linear path of points in $\mathbf{R}^n_+$ starting in some arbitrarily chosen point $z^0 \in \mathbf{R}^n_+$. We allow z^0 to lie on the boundary of $\mathbf{R}^n_+$. In case $z^0 = 0$ the algorithm will coincide with the Lemke complementary pivoting algorithm described in Chapter 4. To motivate the algorithm, let (w^*, z^*) be a solution to $LCP(q, M)$ defined in (4.1). Then it holds that

$$\begin{aligned} \min_h \ w_h^* \geq 0 \quad &\text{if} \quad z^* = 0 \\ \min_h \ w_h^* = 0 \quad &\text{if} \quad z^* \geq 0 \end{aligned}$$

and

$$\begin{aligned} w_j^* = \min_h \ w_h^* \quad &\text{when} \quad z_j^* > 0 \\ w_j^* \geq \min_h \ w_h^* \quad &\text{when} \quad z_j^* = 0 \end{aligned}$$

for all $j \in \mathcal{I}_n$. Let $w^0 = Mz^0 + q$. Then this interpretation of the conditions to hold at a solution (w^*, z^*) of $LCP(q, M)$ makes it very natural to decrease z_j proportionally from z_j^0 for those $j \in \mathcal{I}_n$ for which $w_j^0 > \min\{\min_h w_h^0, 0\}$ and to increase z_j from z_j^0 when $w_j^0 = \min_h w_h^0 < 0$. More precisely, when $w_j^0 = \min_h w_h^0 < 0$ it is natural to leave z^0 towards a point $ae(j)$ for some $a > z_j^0$ on the j-th axis, thereby increasing z_j from z_j^0 and decreasing z_h for $h \neq j$ proportionally from z_h^0. When $\min_h w_h^0 \geq 0$ then it is most natural to leave z^0 towards the origin. This interpretation gives $n+1$ rays to leave the starting point, namely $q(j) = ae(j) - z^0$ for $j \in \mathcal{I}_n$ and $q(n+1) = -z^0$. Given these rays to leave z^0, the algorithm is such that each point z on its path is a stationary point of the function g on the set $\mathcal{H}(t) \cap \mathbf{R}^n_+$ for some $t \geq 0$, where

$$\mathcal{H}(t) = \{z^0 + \textstyle\sum_{j=1}^{n+1} \lambda_j q(j) \ \mid \lambda_j \geq 0 \text{ for } j \in \mathcal{I}_{n+1}, \text{ and } \sum_{j=1}^{n+1} \lambda_j \leq t\}.$$

The number a is an arbitrarily chosen number satisfying $a > \sum_{h=1}^n z_h^0$, assuring at least that $z^0 \in \mathcal{H}(t)$ for $t > 0$.[1]

The number t can be considered as a homotopy parameter running from zero to infinity. For $t = 0$ the set $\mathcal{H}(0) \cap \mathbf{R}^n_+$ only consists of the starting point z^0. Hence z^0 is a stationary point of g on $\mathcal{H}(0) \cap \mathbf{R}^n_+$. For $t = 1$ the set $\mathcal{H}(1) \cap \mathbf{R}^n_+$ is the convex hull of the origin and the points $ae(j)$, $j \in \mathcal{I}_n$, on the axes of $\mathbf{R}^n_+$. If the algorithm generates

[1] In Section 4 of this chapter we will make use of this freedom by letting the choice of a depend on the matrix M and vector q (see Theorem 5.3.1).

a stationary point z of g on $\mathcal{H}(1) \cap \mathrm{bd}\mathbf{R}^n_+$ such that $e^T z < a$ then z is also a stationary point of g on $\mathbf{R}^n_+$ and consequently a solution to the linear complementarity problem. For $t \geq 1$ the set $\mathcal{H}(t) \cap \mathbf{R}^n_+$ is equal to the convex hull of the origin and the points $[(1-t)\sum_{h=1}^n z^0_h + ta]e(j)$, $j \in \mathcal{I}_n$, on the axes of $\mathbf{R}^n_+$.[2] Clearly, the set $\mathcal{H}(t)$ is equal to

$$\mathcal{H}(t) = \{z \in \mathbf{R}^n \mid z \geq (1-t)z^0,\ e^T z \leq (1-t)e^T z^0 + ta\}$$

and hence $\mathcal{H}(t) \cap \mathbf{R}^n_+$ is equal to the set

$$\{z \in \mathbf{R}^n \mid z \geq \max\{1-t, 0\}z^0,\ e^T z \leq (1-t)e^T z^0 + ta\}.$$

For an illustration of the set $\mathcal{H}(t) \cap \mathbf{R}^n_+$ we refer to Figure 5.1.1. Now the algorithm follows a path of stationary points of g on $\mathcal{H}(t) \cap \mathbf{R}^n_+$ for varying parameter $t \geq 0$ starting for $t = 0$ in the arbitrarily chosen point $z^0 \in \mathbf{R}^n_+$ and, barring degeneracy, terminates either on a ray or at a solution.

If $\bar{z}$ in $\mathbf{R}^n_+$ is a stationary point of g on $\mathcal{H}(\bar{t}) \cap \mathbf{R}^n_+$ for some $\bar{t} \geq 0$ then $\bar{z}$ maximizes $z^T g(\bar{z})$ over $\mathcal{H}(\bar{t}) \cap \mathbf{R}^n_+$. By definition of $\mathcal{H}(t)$ this implies that $\bar{z}$ solves the maximization problem, denoted as the primal, given by

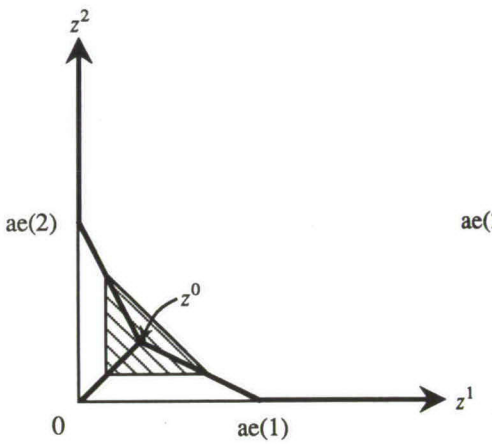
$$\begin{aligned} \max\ & z^T g(\bar{z}) \\ \text{s.t.}\ & z \geq \max\{1-\bar{t}, 0\}z^0 \\ & e^T z \leq (1-\bar{t})e^T z^0 + \bar{t}a. \end{aligned}$$

This maximization problem is a linear programming problem. According to the Duality Theorem of Linear Programming this maximization problem is equivalent to the minimization problem, denoted as the dual, given by

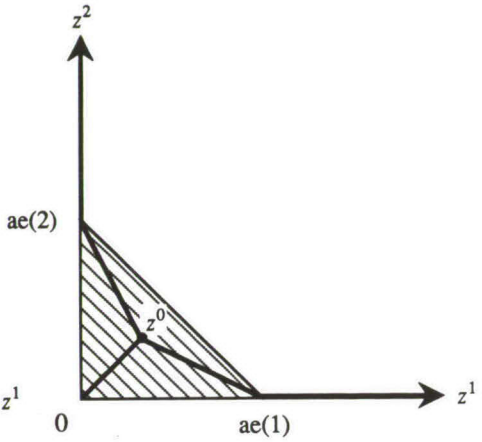
$$\begin{aligned} \min\ & \theta((1-\bar{t})e^T z^0 + \bar{t}a) - \max\{1-\bar{t}, 0\}\mu^T z^0 \\ \text{s.t.}\ & g(\bar{z}) = -\mu + \theta e \\ & \theta \geq 0,\ \mu \geq 0 \end{aligned}$$

where θ is the dual variable to the constraint $e^T z \leq (1-\bar{t})e^T z^0 + \bar{t}a$ and μ the n-vector with μ_j the dual variable to the constraint $z_j \geq \max\{1-\bar{t}, 0\}z^0_j$ for $j \in \mathcal{I}_n$.

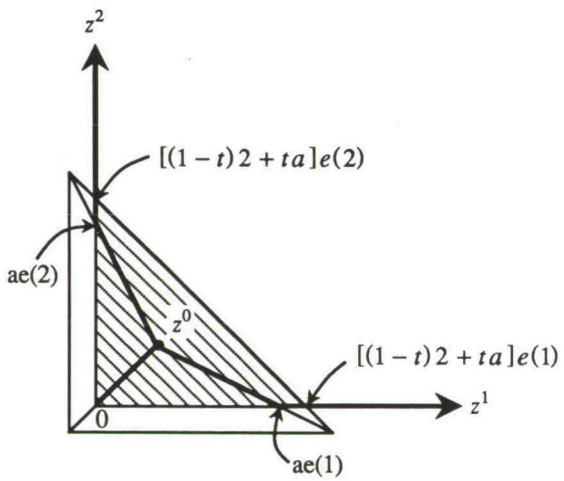
[2] Notice that a point on the k-th axis of $\mathbf{R}^n_+$ is given by $\gamma e(k)$ for some $\gamma \geq 0$. Hence the point of $\mathrm{bd}\mathcal{H}(t)$, $t \geq 1$, on the k-th axis of $\mathbf{R}^n_+$ follows from finding the values of γ, $\lambda_1, \ldots, \lambda_n$ such that $\sum_{j=1}^n \lambda_j = t$ and $\gamma e(k) = z^0 + \sum_{j=1}^{n+1} \lambda_j q(j)$. Adding up the latter equation over all components gives $\gamma = (1-t)\sum_{h=1}^n z^0_h + ta$.



(a) $\mathcal{H}(\frac{1}{2}) \cap \mathbf{R}^2_+$



(b) $\mathcal{H}(1) \cap \mathbf{R}^2_+$



(c) $\mathcal{H}(\frac{3}{2}) \cap \mathbf{R}^2_+$

FIGURE 5.1.1: The subset $\mathcal{H}(t) \cap \mathbf{R}^2_+$ for $t = \frac{1}{2},\ 1,\ \frac{3}{2}$, given $z^0 = (1,1)^\mathsf{T}$ and $a = 3$.

Given $g(\overline{z})$ the dual has a unique solution $\overline{\theta} = \max\{0, \max_h g_h(\overline{z})\}$, $\overline{\mu}_j = \overline{\theta} - g_j(\overline{z})$ for $j \in \mathcal{I}_n$. Let $\overline{\mathcal{T}}$ be a subset of $\mathcal{I}_{n+1}$ such that $\overline{\mu}_j = 0$ for all $j \in \overline{\mathcal{T}} \setminus \{n+1\}$, $\overline{\mu}_j > 0$ for all $j \notin \overline{\mathcal{T}} \cup \{n+1\}$, $\overline{\theta} = 0$ if $n+1 \in \overline{\mathcal{T}}$ and $\overline{\theta} > 0$ if $n+1 \notin \overline{\mathcal{T}}$. Then $\overline{\mu}_j = 0$ and $\overline{z}_j \geq \max\{1-\overline{t}, 0\} z_j^0$ if $j \in \overline{\mathcal{T}} \setminus \{n+1\}$, $\overline{\mu}_j > 0$ and $\overline{z}_j = \max\{1-\overline{t}, 0\} z_j^0$ if $j \notin \overline{\mathcal{T}} \cup \{n+1\}$, $\overline{\theta} = 0$ and $e^T \overline{z} \leq (1-\overline{t}) e^T z^0 + \overline{t} a$ if $n+1 \in \overline{\mathcal{T}}$, and $\overline{\theta} > 0$ and $e^T \overline{z} = (1-\overline{t}) e^T z^0 + \overline{t} a$ if $n+1 \notin \overline{\mathcal{T}}$. Let $\mathcal{T}$-completeness and subsets $A(\mathcal{T})$ and $A^0(\mathcal{T})$ for subsets $\mathcal{T}$ of $\mathcal{I}_{n+1}$ be defined as follows.

Definition 5.1.1 *For a subset $\mathcal{T}$ of $\mathcal{I}_{n+1}$ a point $z \in \mathbf{R}^n_+$ is $\mathcal{T}$-complete if $j \in \mathcal{T}$ implies $g_j(\overline{z}) = \theta$ and $n+1 \in \mathcal{T}$ implies $\theta = 0$, where $\theta = \max\{\max_h g_h(z), 0\}$.*

Definition 5.1.2 *For $\mathcal{T} \subset \mathcal{I}_{n+1}$*

$$A(\mathcal{T}) = \emptyset \text{ if } n+1 \in T \text{ and } z_h^0 = 0 \text{ for all } h \notin T$$

and otherwise

$$A(\mathcal{T}) = (\{z^0\} + \text{cone}(\{q(j) \mid j \in T\})) \cap \mathbf{R}^n_+.$$

Definition 5.1.3 *For $\mathcal{T} \subset \mathcal{I}_{n+1}$*

$$A^0(\mathcal{T}) = \emptyset \text{ if } n+1 \in T \text{ or } z_h^0 = 0 \text{ for all } h \notin T$$

and otherwise

$$A^0(\mathcal{T}) = \{\sum_{j \in T} \lambda_j a e(j) \mid \lambda_j \geq 0 \text{ for } j \in T \text{ and } \sum_{j \in T} \lambda_j \geq 1\}.$$

Figure 5.1.2 gives a subdivision of $\mathbf{R}^n_+$ into subsets $A(\mathcal{T})$ and $A^0(\mathcal{T})$ for subsets $\mathcal{T}$ of $\mathcal{I}_{n+1}$ when $n = 2$.

Theorem 5.1.1 *The point $z \in \mathbf{R}^n_+$ is a $\mathcal{T}$-complete point in $A(\mathcal{T})$ or $A^0(\mathcal{T})$ for some $\mathcal{T} \subset \mathcal{I}_{n+1}$, if and only if z is a stationary point of g on $\mathcal{H}(t) \cap \mathbf{R}^n_+$ for some $t \geq 0$.*

Proof: Similar to the proof of Theorem 4.2.1.

To generate a path of stationary points of g on $\mathcal{H}(t) \cap \mathbf{R}^n_+$ from z^0 for varying $t \geq 0$ the algorithm could therefore generate $\mathcal{T}$-complete points in $A(\mathcal{T})$ or $\mathcal{T}$-complete points in $A^0(\mathcal{T})$ for varying subsets $\mathcal{T}$ of $\mathcal{I}_{n+1}$. The next section describes how to generate this path through subsequent subsets $A(\mathcal{T})$ and $A^0(\mathcal{T})$ for varying subsets $\mathcal{T}$ of $\mathcal{I}_{n+1}$.

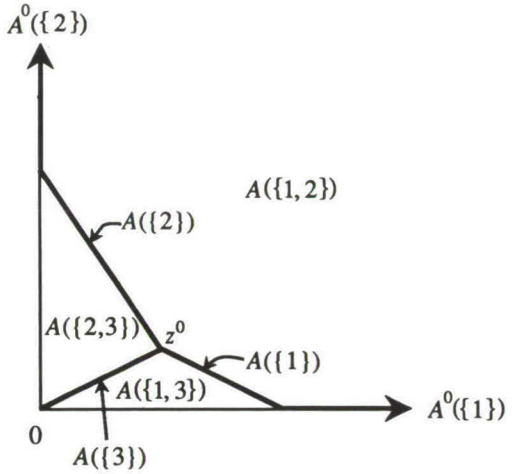


FIGURE 5.1.2: Subdivision of $\mathbf{R}^2_+$ into subsets $A(\mathcal{T})$ and $A^0(\mathcal{T})$ for $\mathcal{T} \subset \mathcal{I}_3$.

5.2 The steps of the algorithm

Definition 5.1.1 leads to a pivot system in each point z on the path generated by the algorithm from the starting point z^0 either towards a solution of the linear complementarity problem or towards infinity. The $\mathcal{T}$-completeness condition at a point z is equivalent to the existence of $\mu_j \geq 0$ $(j \notin \mathcal{T} \cup \{n+1\})$, $\theta \geq 0$ if $n+1 \notin \mathcal{T}$ and $\theta = 0$ if $n+1 \in \mathcal{T}$ such that

$$-Mz - q = \theta e - \sum_{j \notin \mathcal{T} \cup \{n+1\}} \mu_j e(j).$$

Combined with $z \in A(\mathcal{T})$ or $z \in A^0(\mathcal{T})$ the appropriate pivot system for $\mathcal{T}$-completeness at a point z in $A(\mathcal{T})$ or $A^0(\mathcal{T})$ is given in one of the next two lemma's.

Lemma 5.2.1 *A point $z \in A(\mathcal{T})$ is $\mathcal{T}$-complete for some feasible $\mathcal{T} \subset \mathcal{I}_{n+1}$ if and only if the system of equations*

$$\sum_{j \in \mathcal{T}} \lambda_j M q(j) - \sum_{j \notin \mathcal{T} \cup \{n+1\}} \mu_j e(j) + \theta e = -q - Mz^0 \tag{5.1}$$

has a solution $\lambda_j \geq 0$ $(j \in \mathcal{T})$, $\mu_j \geq 0$ $(j \notin \mathcal{T} \cup \{n+1\})$, $\theta \geq 0$ if $n+1 \notin \mathcal{T}$ and $\theta = 0$ if $n+1 \in \mathcal{T}$, such that $z = z^0 + \sum_{j \in \mathcal{T}} \lambda_j q(j)$.

Lemma 5.2.2 *A point $z \in A(\mathcal{T})$ for some $\mathcal{T} \subseteq \mathcal{I}_n$ with $z_i^0 = 0$ for all $i \notin \mathcal{T}$ or a point $z \in A^0(\mathcal{T})$ for some $\mathcal{T} \subseteq \mathcal{I}_n$ is $\mathcal{T}$-complete if and only if the system of equations*

$$\sum_{j \in \mathcal{T}} \lambda_j a M_{.j} - \sum_{j \notin \mathcal{T} \cup \{n+1\}} \mu_j e(j) + \theta e = -q \tag{5.2}$$

has a solution $\lambda_j \geq 0$ $(j \in \mathcal{T})$, $\mu_j \geq 0$ $(j \notin \mathcal{T} \cup \{n+1\})$, and $\theta \geq 0$, such that $z = \sum_{j \in \mathcal{T}} \lambda_j a_j e(j)$.

Notice that the pivot systems in (5.1) and (5.2) both contain n equations in $n+1$ variables leaving us with one degree of freedom. If nonempty, the solution set of each system forms a line segment, assuming nondegeneracy. This line segment corresponds to a linear piece of $\mathcal{T}$-complete points in $A(\mathcal{T})$ or in $A^0(\mathcal{T})$ with either one or two end points. As we will show below each end point of a line segment of solutions to a system of equations for some $\mathcal{T} \subset \mathcal{I}_{n+1}$ either corresponds to the starting point z^0 or to a solution to the linear complementarity problem or is an end point of a line segment of solutions to exactly one other system of equations possibly for a different set $\mathcal{T}$. The point z^0 corresponds to an end point of only one line segment of solutions. These properties make the path of points generated by the algorithm from z^0 a piecewise linear path through subsequent subsets $A(\mathcal{T})$ and $A^0(\mathcal{T})$ for varying subsets $\mathcal{T}$ of $\mathcal{I}_{n+1}$. Each linear piece can be followed by making a linear programming pivot step in the appropriate pivot system with the variable being zero (or making a binding constraint) at an end point.

A linear piece of $\mathcal{T}$-complete points in $A(\mathcal{T})$ for some subset $\mathcal{T} \subseteq \mathcal{I}_n$ for which $z_i^0 = 0$ for all $i \notin \mathcal{T}$ can be generated by making a pivot step in system (5.1) or in system (5.2). Which one of these systems will be appropriate depends on in which system the previous pivoting step was made. This feature causes the algorithm to generate the path through different subsets of $\mathbf{R}_+^n$ in an efficient way. Changing from one pivot system to the other one at an end point of a line segment requires a redefinition of the variables λ_j, $j \in T$. The setup in Lemma 5.2.1 and Lemma 5.2.2 allows us to make as few of these changes of variables as possible.

Suppose the algorithm is following a linear piece of $\mathcal{T}$-complete points in $A(\mathcal{T})$ or in $A^0(\mathcal{T})$ for some $\mathcal{T} \subset \mathcal{I}_{n+1}$, i.e., a pivot step is made in one of the systems of equations (5.1) or (5.2) with the variable being zero at an end point of the line segment of solutions. When the linear piece has another end point, say $\bar{z}$, then,

assuming nondegeneracy, exactly one of the following cases will occur for the solution $(\overline{\lambda}, \overline{\mu}, \overline{\beta})$ at this end point:

Case 1: $\overline{\lambda}_p$ is equal to zero for some $p \in \mathcal{T}$, while $\mathcal{T} \setminus \{p\} \neq \emptyset$. Then $\overline{z}$ is an end point lying in $A(\mathcal{T} \setminus \{p\})$ or in $A^0(\mathcal{T} \setminus \{p\})$ depending on whether $\overline{z} \in A(\mathcal{T})$ or $\overline{z} \in A^0(\mathcal{T})$ respectively. The algorithm proceeds in $A(\mathcal{T} \setminus \{p\})$ or $A^0(\mathcal{T} \setminus \{p\})$ by pivoting the column $-e(p)$ into the appropriate system of equations thereby raising μ_p from zero if $p \neq n+1$, and the column e thereby raising θ from zero if $p = n+1$ in order to maintain $\mathcal{T} \setminus \{p\}$-completeness.

Case 2: In system (5.1) $\overline{\lambda}_p$ is equal to

$$(\sum_{j \in \mathcal{T} \setminus \{p\}} \overline{\lambda}_j - 1)\left(\frac{z_p^0}{a - z_p^0}\right)$$

for some $p \in \mathcal{T}$. Then $\overline{z}$ is an end point lying in $A^0(\mathcal{T} \setminus \{p\})$. Let

$$\hat{\lambda}_j = \overline{\lambda}_j + (1 - \sum_{h \in \mathcal{T}} \overline{\lambda}_h)\left(\frac{z_j^0}{a}\right)$$

for $j \in \mathcal{T} \setminus \{p\}$. Then $\hat{\lambda}_j \geq 0$ $(j \in \mathcal{T} \setminus \{p\})$, $\overline{\mu}_h \geq 0$ $(h \notin \mathcal{T} \cup \{n+1\})$, $\overline{\mu}_p = 0$, and $\overline{\theta} \geq 0$ is a solution to system (5.2) and $\overline{z}$ is an end point of a linear piece of $\mathcal{T} \setminus \{p\}$-complete points in $A^0(\mathcal{T} \setminus \{p\})$. The algorithm proceeds in $A^0(\mathcal{T} \setminus \{p\})$ by changing system (5.1) into system (5.2) and pivoting the column $-e(p)$ into the new system (5.2) thereby raising μ_p from zero in order to maintain $\mathcal{T} \setminus \{p\}$-completeness.

Case 3: $\sum_{j \in \mathcal{T}} \overline{\lambda}_j$ is equal to 1 in system (5.1) while $n+1 \in \mathcal{T}$ or $z_h^0 > 0$ for some $h \notin \mathcal{T}$, or in system (5.2) while $z_h^0 > 0$ for some $h \notin \mathcal{T}$. Suppose $n+1 \in \mathcal{T}$ then $\overline{w}_j = 0$ and $\overline{z}_j = \overline{\lambda}_j a \geq 0$ for $j \in \mathcal{T}$ while $\overline{w}_j = \overline{\mu}_j \geq 0$ and $\overline{z}_j = (1 - \sum_{j \in \mathcal{T}} \overline{\lambda}_j) z_j^0 = 0$ for $j \notin \mathcal{T}$, leaving us with a solution to the linear complementarity problem in $\overline{z}$. Otherwise, suppose $n+1 \notin \mathcal{T}$. Then $\overline{z}$ is an end point of a linear piece of $\mathcal{T}$-complete points in $A(\mathcal{T})$ as well as in $A^0(\mathcal{T})$. So, if $\overline{z}$ were the end point of a linear piece of $\mathcal{T}$-complete points in $A(\mathcal{T})$ then the algorithm proceeds by generating a linear piece of $\mathcal{T}$-complete points in $A^0(\mathcal{T})$. This linear piece of $\mathcal{T}$-complete points in $A^0(\mathcal{T})$ is generated by changing system (5.1) into system (5.2) and pivoting the column $aM_{\cdot k}$ or $-e(k)$ into the new system (5.2), depending on whether $Mq(k)$ or $-e(k)$ was the last pivot column in (5.1). Notice that $\sum_{j \in \mathcal{T}} \lambda_j$ is then raised from 1. Conversely, if $\overline{z}$ were the end point of a linear piece of $\mathcal{T}$-complete points in $A^0(\mathcal{T})$ then the algorithm

proceeds by generating a linear piece of $\mathcal{T}$-complete points in $A(\mathcal{T})$. This linear piece of $\mathcal{T}$-complete points in $A(\mathcal{T})$ is generated by changing system (5.2) into system (5.1) and pivoting the column $Mq(k)$ or $-e(k)$ into the new system (5.1), depending on whether $aM_{.k}$ or $-e(k)$ was the last pivot column in the system (5.2). Hence $\sum_{j\in\mathcal{T}} \lambda_j$ is lowered from 1.

Case 4: In system (5.2) it holds that for some $p \in \mathcal{T}$

$$\sum_{i\in\mathcal{T}\setminus\{p\}} \overline{\lambda}_i z_p^0 + \overline{\lambda}_p \left(a - \sum_{i\in\mathcal{T}} z_i^0\right) = z_p^0 \text{ while } z_p^0 > 0.$$

Then $\overline{z}$ is an end point lying in $A(\mathcal{T} \setminus \{p\})$. Let

$$\hat{\lambda}_h = \overline{\lambda}_h + z_h^0 \left(\frac{1 - \sum_{j\in\mathcal{T}\setminus\{p\}} \overline{\lambda}_j}{a - \sum_{j\in\mathcal{T}} z_j^0}\right)$$

for $h \in \mathcal{T} \setminus \{p\}$. Then $\hat{\lambda}_h \geq 0$ $(h \in \mathcal{T} \setminus \{p\})$, $\overline{\mu}_h \geq 0$ $(h \notin \mathcal{T} \cup \{n+1\})$, $\overline{\mu}_p = 0$, and $\overline{\theta} \geq 0$ is a solution to (5.1) and $\overline{z}$ is an end point of a linear piece of $\mathcal{T} \setminus \{p\}$-complete points in $A(\mathcal{T} \setminus \{p\})$. The algorithm proceeds by changing system (5.2) into system (5.1) and pivoting the column $-e(p)$ into system (5.1) thereby raising μ_p from zero in order to maintain $\mathcal{T} \setminus \{p\}$-completeness.

Case 5: $\overline{\mu}_k$ is zero for some $k \notin \mathcal{T} \cup \{n+1\}$. Suppose $\overline{z} \in A(\mathcal{T})$. If $z_h^0 = 0$ for all $h \notin \mathcal{T} \cup \{k\}$ while $n+1 \in \mathcal{T}$ or if $\mathcal{T} \cup \{k\} = \mathcal{I}_{n+1}$ then $\overline{z}$ is a solution to the linear complementarity problem. Otherwise $\overline{z}$ is an end point of a linear piece of $\mathcal{T} \cup \{k\}$-complete points in $A(\mathcal{T} \cup \{k\})$. The algorithm proceeds by pivoting the column $Mq(k)$ into the system (5.1) or $aM_{.k}$ into the system (5.2) thereby raising λ_k from zero in order to maintain $\mathcal{T} \cup \{k\}$-completeness.

Suppose $\overline{z} \in A^0(\mathcal{T})$. If $z_h^0 = 0$ for all $h \notin \mathcal{T} \cup \{k\}$ then $\overline{z}$ is an end point of a linear piece of $\mathcal{T} \cup \{k\}$-complete points in $A(\mathcal{T} \cup \{k\})$, otherwise $\overline{z}$ is an end point of a linear piece of $\mathcal{T} \cup \{k\}$-complete points in $A^0(\mathcal{T} \cup \{k\})$. The algorithm proceeds in both cases by pivoting the column $aM_{.k}$ into the system (5.2) thereby raising λ_k from zero in order to maintain $\mathcal{T} \cup \{k\}$-completeness.

Case 6: $\overline{\theta}$ is zero. Then $\overline{z}$ is a solution to the linear complementarity problem if $\overline{z} \in A^0(\mathcal{T})$ or if $\overline{z} \in A(\mathcal{T})$ and $z_h^0 = 0$ for all $h \notin \mathcal{T}$. Otherwise, $\overline{z}$ is an end point of a linear piece of $\mathcal{T} \cup \{n+1\}$-complete points in $A(\mathcal{T} \cup \{n+1\})$. The algorithm proceeds by pivoting the column $-Mz^0$ into (5.1) thereby raising λ_{n+1} from zero in order to maintain $\mathcal{T} \cup \{n+1\}$-completeness.

The cases 1 to 6 describe the performance of the algorithm at the end points of all possible line segments generated by the algorithm except at z^0 where the algorithm is initiated. To show that z^0 is an end point of a (unique) linear piece of $\mathcal{T}$-complete points in $A(\mathcal{T})$ for some subset $\mathcal{T}$ of $\mathcal{I}_{n+1}$ let us denote Mz^0+q by w^0. If $\min_h w_h^0 < 0$ let k be such that $w_k^0 = \min_h w_h^0$. Then the starting point z^0 is $\mathcal{T}^0$-complete with $\mathcal{T}^0 = \{k\}$ and the system of equations

$$-\sum_{j\neq k,n+1} \mu_j e(j) + \theta e = -q - Mz^0 \tag{5.3}$$

has a unique solution $\mu_j^0 = w_j^0 - w_k^0 > 0$ $(j \neq k, n+1)$, and $\theta^0 = -w_k^0 > 0$. So, assuming nondegeneracy, z^0 is an end point of a linear piece of $\{k\}$-complete points in $A(\{k\})$. In order to follow this linear piece the algorithm starts by pivoting the column $Mq(k)$ into (5.3) thereby raising λ_k from zero.

If $\min_h w_h^0 \geq 0$ then the starting point z^0 solves the linear complementarity problem if $z^0 = 0$. Otherwise z^0 is $\mathcal{T}^0$-complete with $\mathcal{T}^0 = \{n+1\}$ and the system of equations

$$-\sum_{j=1}^{n} \mu_j e(j) = -q - Mz^0 \tag{5.4}$$

has a unique solution $\mu_j^0 = w_j^0 > 0$ $(j \in \mathcal{I}_n)$. Assuming nondegeneracy, z^0 is the end point of a linear piece of $\{n+1\}$-complete points in $A(\{n+1\})$. In order to follow this linear piece the algorithm starts by pivoting the column $-Mz^0$ into (5.4) thereby raising λ_{n+1} from zero.

5.3 Convergence issues

Starting in some arbitrarily chosen point $z^0 \in \mathbf{R}_+^n$ the algorithm generates a piecewise linear path of $\mathcal{T}$-complete points through adjacent subsets $A(\mathcal{T})$ or $A^0(\mathcal{T})$, for varying subsets $\mathcal{T}$ of $\mathcal{I}_{n+1}$, as described in Section 2. This path either ends up with a solution to the linear complementarity problem as defined in (4.1) or it ends up with a ray towards infinity. The end points giving rise to a solution to the linear complementarity problem have already been described during the enumeration of the cases in Section 2. Lemma 5.3.1 summarizes all the cases in which the algorithm ends up with a solution.

Lemma 5.3.1 *Let z be an end point of a linear piece of $\mathcal{T}$-complete points on the path generated by the algorithm in $A(\mathcal{T})$ or in $A^0(\mathcal{T})$ for some $\mathcal{T} \subset \mathcal{I}_{n+1}$. Then z is a solution to the linear complementarity problem if one of the following cases holds:*

i) $z \in A(\mathcal{T})$, $n+1 \in T$, $\mu_k = 0$ *for some* $k \notin T$, *and* $z_h^0 = 0$ *for all* $h \notin \mathcal{T} \cup \{k\}$ *or* $\mathcal{T} \cup \{k\} = \mathcal{I}_{n+1}$;

ii) $z \in A^0(\mathcal{T})$ *and* $\theta = 0$;

iii) $z \in A(\mathcal{T})$, $z_h^0 = 0$ *for all* $h \notin T$, *and* $\theta = 0$;

iv) $z \in A(\mathcal{T})$, $n+1 \in T$, *and* $\sum_{j \in T} \lambda_j = 1$,

where the variables $\lambda_j \geq 0$ $(j \in T)$, $\mu_j \geq 0$ $(j \notin \mathcal{T} \cup \{n+1\})$, *and* $\theta \geq 0$ *are the solution to the appropriate pivot system at* z.

The possibility of divergence urges us to impose a convergence condition on the problem. Notice that divergence can only occur when the algorithm is generating a path of points in $A^0(\mathcal{T})$ or in $A(\mathcal{T})$ with $\mathcal{T}$ such that $z_i^0 = 0$ for all $i \notin \mathcal{T}$ and $n+1 \notin \mathcal{T}$, i.e., when the system of equations in Lemma 5.2.2 is appropriate. Therefore we can restrict our attention to the possible occurence of a ray of solutions to system (5.2) for some $\mathcal{T} \subseteq \mathcal{I}_n$. System (5.2) however is equivalent to (4.8), the system used in Lemke's algorithm to solve the linear complementarity problem. So, the convergence theorem on Lemke's algorithm given in Theorem 4.2.3 can be used for our algorithm.

What remains is a more precise delimitation of the possibilities of choosing the number a. In case $z^0 = 0$, the algorithm coincides with Lemke's algorithm. In this case pivot steps need only to be made in system (5.2) and the number a can be set equal to one. Suppose now that $z^0 \neq 0$. In that case we already put one limitation on a being independent of the problem as defined in (4.1) but guaranteeing that each $A(\mathcal{T})$, $\mathcal{T} \subset \mathcal{I}_{n+1}$, is convex. To make the choice of a dependent on the data of the problem, i.e., on M and q, we suggest to choose a such that for all $j \in \mathcal{I}_n$ no $\{j\}$-complete points in $A^0(\{j\})$ can be found. For every j, let a be such that no j-complete points in $A^0(\{j\})$ exist. This implies that the system (5.2) for $\mathcal{T}$ equal to $\{j\}$,

$$\lambda_j a M_{.j} - \sum_{h \neq j, n+1} \mu_h e(h) + \theta e = -q, \tag{5.5}$$

may not have a solution $\lambda_j \geq 1$, $\mu_h \geq 0$ $(h \neq j, n+1)$, $\theta \geq 0$. The following condition on a assures that at a solution to (5.5) for $\lambda_j > 1$ it holds that $\theta < 0$ or $\mu_h < 0$ for some $h \neq j, n+1$:

$$\begin{aligned} &\text{if } M_{jj} > 0 \text{ then } a > \min\left\{\tfrac{-q_j}{M_{jj}},\ \min_{h:M_{hj}<M_{jj}}\left\{\tfrac{q_h-q_j}{M_{jj}-M_{hj}}\right\}\right\}, \\ &\text{if } M_{jj} = 0 \text{ then } a > \min_{h:M_{hj}<M_{jj}}\left\{\tfrac{q_j-q_h}{M_{hj}}\right\}. \end{aligned} \tag{5.6}$$

Of course we assume M to fulfil the conditions imposed by the convergence theorem, Theorem 4.2.3. Then $M_{jj} \geq 0$ for all $j \in \mathcal{I}_n$. If these conditions do not hold it is possible that for some $j \in \mathcal{I}_n$ a can not be calculated according to (5.6). Then one knows in advance that the algorithm could diverge and that the linear complementarity problem might not even have a solution.

Condition (5.6) suggests how to determine the number a.

Theorem 5.3.1 *Suppose M can be written as $P + C$ where P is copositive-plus and symmetric and C is copositive, and the system $q + Px - C^T y \geq 0$, $y \geq 0$ is feasible. Let a be chosen such that $a > \max\{\sum_{h=1}^{n} z_h^0, a_1, \ldots, a_n\}$ where a_j is such that*

$$\textit{if } M_{jj} > 0 \textit{ then } a_j > \min\left\{\tfrac{-q_j}{M_{jj}},\ \min_{h:M_{hj}<M_{jj}}\left\{\tfrac{q_h-q_j}{M_{jj}-M_{hj}}\right\}\right\},$$

$$\textit{if } M_{jj} = 0 \textit{ then } a_j > \min_{h:M_{hj}<M_{jj}}\left\{\tfrac{q_j-q_h}{M_{hj}}\right\},$$

for all $j \in \mathcal{I}_n$. Then the algorithm always converges and can not generate $\{j\}$-complete points in $A^0(\{j\})$, $j \in \mathcal{I}_{n+1}$.

Chapter 6

The generalized nonlinear complementarity problem

The *generalized nonlinear complementarity* problem extends the nonlinear complementarity problem introduced in Chapter 2 to the incorporation of arbitrary upper and lower bounds on the variables. The problem is therefore also referred to as the nonlinear complementarity problem with lower and upper bounds. It is defined as follows.

> Given two vectors a and b in $\mathbf{R}^n$ with $a_i < b_i$ for all $i \in \mathcal{I}_n$ and a continuous function $f : \mathbf{R}^n \to \mathbf{R}^n$, find an $x^* \in \mathbf{R}^n$ such that $a \leq x^* \leq b$ and for all $i \in \mathcal{I}_n$
>
> $$\begin{aligned} f_i(x^*) &\leq 0 \text{ if } a_i = x_i^* \\ f_i(x^*) &= 0 \text{ if } a_i < x_i^* < b_i \\ f_i(x^*) &\geq 0 \text{ if } x_i^* = b_i. \end{aligned} \tag{6.1}$$

We allow components of the vector a to be $-\infty$ and of the vector b to be $+\infty$. The generalized nonlinear complementarity problem encloses many well-known problems under which all the complementarity problems introduced in the previous chapters as well as the linearized version of (6.1) called the generalized linear complementarity problem. An algorithm to solve the generalized linear complementarity problem has been introduced in van der Laan and Talman (1985). The algorithm introduced in this chapter can be seen as its nonlinear counterpart and follows ideas of Talman and Yamamoto (1989) for solving the stationary point problem on a polytope. Further-

more our algorithm is a natural alternative to the simplicial algorithm introduced in van der Laan and Talman (1987) to solve the generalized nonlinear complementarity problem.

In this chapter we introduce a simplicial variable dimension restart algorithm. The algorithm subdivides the set on which the problem is defined into simplices and generates from an arbitrarily chosen starting point a piecewise linear path of points leading to an approximating solution. When the accuracy is not sufficient the algorithm can be restarted at the approximating solution with a finer simplicial subdivision. The path of points generated by the algorithm can be followed by a sequence of adjacent simplices of varying dimension and can be interpreted as a path of stationary points of the piecewise linear approximation of f on a set with parametrized bounds.

This chapter is based on Kremers and Talman (1990a) and is subdivided in the following way. In Section 1 of this chapter we introduce the path which will approximately be followed by the algorithm. The steps to generate a linear piece of the piecewise linear path followed by the algorithm are described in Section 2. Section 3 introduces an appropriate simplicial subdivision of $\mathcal{C}^n$ to underly the algorithm.

6.1 The path to be approximated by the algorithm

In Theorem 2.3.3 it was shown that all complementarity problems introduced in the previous chapters are equivalent to a stationary point problem on the underlying set $\mathbf{R}^n_+$. This provided a reason to solve these problems by using an algorithm which followed a path of stationary points on a parametrized set. In this section we first show that the nonlinear complementarity problem with lower bound a and upper bound b is equivalent to the stationary point problem of f on the set $\mathcal{C}^n := \{x \in \mathbf{R}^n \mid a \leq x \leq b\}$. In this chapter we assume that all the components of both a and b are finite. The non-finite case is treated in the next chapter. Secondly we describe the path of points in the set $\mathcal{C}^n$ followed approximately by the algorithm. Each point on the path is a stationary point of f on $\mathcal{C}^n$ with parametrized bounds.

Theorem 6.1.1 *Let $f : \mathcal{C}^n \to \mathbf{R}^n$ be a continuous function. Then a point x^* solves the generalized nonlinear complementarity problem if and only if x^* is a stationary point of f on $\mathcal{C}^n$.*

Proof: The point x^* is a stationary point of f on $\mathcal{C}^n$ if and only if x^* solves the maximization problem given by

$$\begin{aligned} &\max\ x^\top f(x^*) \\ &\text{s.t. } a \leq x \leq b. \end{aligned}$$

This problem is a linear programming problem. Hence, according to the Duality Theorem of Linear Programming, solving this problem is equivalent to solving its dual problem given by

$$\begin{aligned} &\min\ b^\top y^2 - a^\top y^1 \\ &\text{s.t. } f(x^*) = y^2 - y^1 \\ &\qquad y^1 \geq 0,\ y^2 \geq 0 \end{aligned}$$

where y^1 is the vector of dual variables for $-x \leq -a$ and y^2 for $x \leq b$. Clearly, since $b > a$ this dual problem coincides with the generalized nonlinear complementarity problem having x^* as a solution. □

Starting in an arbitrarily chosen point $v \in \mathcal{C}^n$ the algorithm follows approximately a path of points x in $\mathcal{C}^n$ such that a point on the path is a stationary point of f on the parametrized set $\mathcal{C}^n_\rho := (1-\rho)\{v\} + \rho\mathcal{C}^n$ for some ρ, $0 \leq \rho \leq 1$. Given the equivalence between the stationary point problem and the generalized nonlinear complementarity problem this implies that at every x on the path for all $i \in \mathcal{I}_n$ it holds that

$$\begin{aligned} &f_i(x) \leq 0 \text{ if } (1-\rho)v_i + \rho a_i = x_i \\ &f_i(x) = 0 \text{ if } (1-\rho)v_i + \rho a_i < x_i < (1-\rho)v_i + \rho b_i \\ &f_i(x) \geq 0 \text{ if } \qquad\qquad x_i = (1-\rho)v_i + \rho b_i \end{aligned} \tag{6.2}$$

for some ρ, $0 \leq \rho \leq 1$. Under some regularity and nondegeneracy conditions, see van den Elzen (1991), the set of points x being a solution to (6.2) for ρ, $0 \leq \rho \leq 1$, constitute piecewise smooth curves in $\mathcal{C}^n$. Each of these curves is either a loop or a path with two end points. One of these paths, say $\mathcal{P}$, has v as an end point for $\rho = 0$. All other end points of paths are solutions to (6.1).

Without loss of generality we may assume that no component of $f(v)$ equals zero. Then, by increasing ρ from 0, the path $\mathcal{P}$ leaves v through increasing x_i from v_i such that $x_i = (1-\rho)v_i + \rho b_i$ if $f_i(v) > 0$ and through decreasing x_i from v_i such that $x_i = (1-\rho)v_i + \rho a_i$ if $f_i(v) < 0$, for all $i \in \mathcal{I}_n$. Hence by increasing ρ from zero the path $\mathcal{P}$ leaves v in the direction pointing towards a corner point of $\mathcal{C}^n$. If along the path $\mathcal{P}$ at a stationary point $x = (1-\rho)v + \rho z$, of f on $\mathcal{C}^n_\rho$ with ρ between 0 and 1 and z a point in the boundary of $\mathcal{C}^n$, $f_j(x)$ becomes zero for some $j \in \mathcal{I}_n$ while $z_j = a_j$ (or b_j), then either x solves (6.1) or the path continues by increasing x_j from $(1-\rho)v_j + \rho a_j$ (decreasing x_j from $(1-\rho)v_j + \rho b_j$) and keeping $f_j(x) = 0$. If at a point x on $\mathcal{P}$, x_j becomes equal to $(1-\rho)v_j + \rho a_j$ (or $(1-\rho)v_j + \rho b_j$) for some $j \in \{i \mid f_i(x) = 0\}$, then the path $\mathcal{P}$ proceeds by decreasing (increasing) $f_j(x)$ from zero and keeping $x_j = (1-\rho)v_j + \rho a_j$ (or $(1-\rho)v_j + \rho b_j$). Finally, if at a point x on $\mathcal{P}$, ρ becomes equal to 1, then, because $\mathcal{C}^n_1 = \mathcal{C}^n$ and hence the conditions in (6.2) reduce to (6.1), the point x is a solution to the generalized nonlinear complementarity problem in (6.1) and thereby an end point of the path $\mathcal{P}$ in $\mathcal{C}^n$. In this way the path $\mathcal{P}$ leads in a very natural way from v to a solution of (6.1). In Figure 6.1.1, $\mathcal{C}^n_\rho$ is drawn for some values of ρ when $n = 2$. Also the rays along which the algorithm may leave v are shown.

6.2 The algorithm

The algorithm approximately follows the path $\mathcal{P}$ described in the previous section by generating a piecewise linear path $\overline{\mathcal{P}}$ connecting v with an approximating solution $\overline{x}$ of (6.1). To describe this piecewise linear path we approximate the function f by its piecewise linear approximation F on a simplicial subdivision of $\mathcal{C}^n$ which we denote by $\mathcal{G}$. For an appropriate simplicial subdivision of $\mathcal{C}^n$ we refer to Section 3 of this chapter.

The results obtained in Section 1 with respect to f can also be applied to the piecewise linear approximation F of f. In particular, there exists a piecewise linear path $\overline{\mathcal{P}}$ of points in $\mathcal{C}^n$ connecting v and a solution to (6.1) with respect to F. For

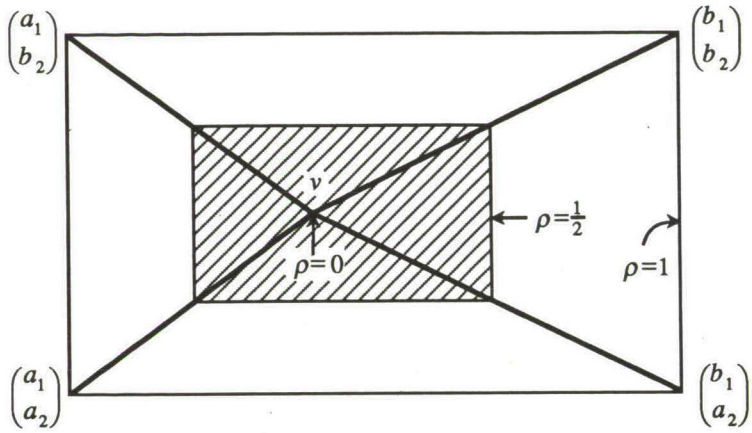


FIGURE 6.1.1: $\mathcal{C}^2_\rho$ for $\rho = 0$, $\frac{1}{2}$, and 1. The linear pieces connecting v with the corner points of $\mathcal{C}^2$ denote the rays along which the path may leave v.

each point x on the path $\overline{\mathcal{P}}$ there exists a ρ between 0 and 1 such that for all $i \in \mathcal{I}_n$

$$\begin{aligned} &F_i(x) \leq 0 \text{ if } (1-\rho)v_i + \rho a_i = x_i \\ &F_i(x) = 0 \text{ if } (1-\rho)v_i + \rho a_i < x_i < (1-\rho)v_i + \rho b_i \\ &F_i(x) \geq 0 \text{ if } \qquad\qquad x_i = (1-\rho)v_i + \rho b_i. \end{aligned} \tag{6.3}$$

This leads to the following definitions.

Definition 6.2.1 *For each sign vector $s \in \mathbf{R}^n$ the face $\mathcal{C}^n(s)$ of $\mathcal{C}^n$ is given by*

$$\begin{aligned} \mathcal{C}^n(s) = \{x \in \mathcal{C}^n \mid\ & x_i = a_i \text{ if } s_i = -1 \text{ for all } i, \text{ and} \\ & x_i = b_i \text{ if } s_i = +1\}. \end{aligned}$$

Given this definition of $\mathcal{C}^n(s)$ for a sign vector $s \in \mathbf{R}^n$ we can define a subset $A(s)$ of $\mathcal{C}^n$ as follows.

Definition 6.2.2 *For each sign vector $s \in \mathbf{R}^n$ the subset $A(s)$ of $\mathcal{C}^n$ is given by*

$$A(s) = \emptyset \text{ if } v \in \mathcal{C}^n(s)$$

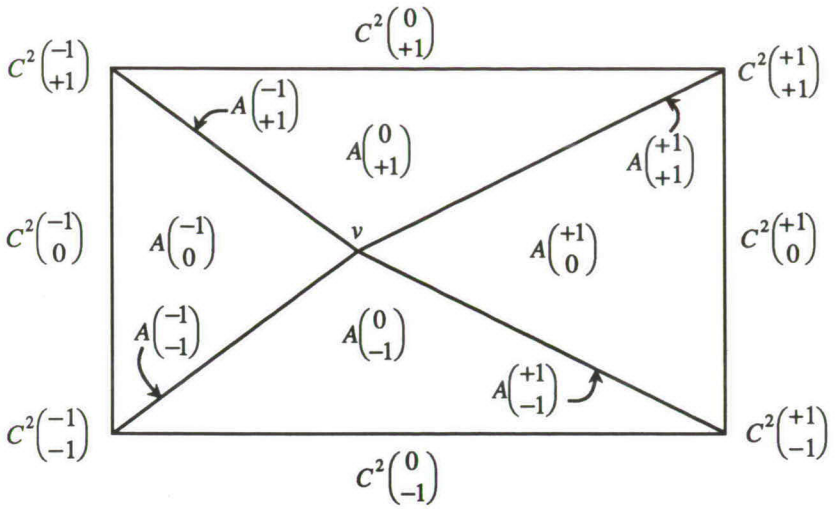


FIGURE 6.2.1: Subdivision of $\mathcal{C}^2$ into subsets $A(s)$ and $\mathcal{C}^2(s)$ for sign vectors $s \in \mathbf{R}^2$.

and otherwise

$$A(s) = \{x \in \mathcal{C}^n \mid \begin{array}{ll} (1-\rho)v_i + \rho a_i = x_i & \text{if } s_i = -1 \\ (1-\rho)v_i + \rho a_i \leq x_i \leq (1-\rho)v_i + \rho b_i & \text{if } s_i = 0 \\ x_i = (1-\rho)v_i + \rho b_i & \text{if } s_i = +1 \\ \text{for } 0 \leq \rho \leq 1\}. & \end{array}$$

Notice that in case v does not lie in the face $\mathcal{C}^n(s)$, $A(s)$ is the convex hull of the point v and $\mathcal{C}^n(s)$. Figure 6.2.1 gives a subdivision of $\mathcal{C}^n$ into subsets $A(s)$ and $\mathcal{C}^n(s)$ for sign vectors $s \in \mathbf{R}^n$ when $n = 2$.

The path $\overline{\mathcal{P}}$ generated by the algorithm is then a path of points $x \in A(s)$ such that $s = \text{sgn}(F(x))$ for varying sign vectors $s \in \mathbf{R}^n$. Moreover, a point $\overline{x}$ on $\overline{\mathcal{P}}$ is a solution to (6.1) with respect to F if $\overline{x} \in \mathcal{C}^n(s)$ and $s = \text{sgn}(F(\overline{x}))$.

The simplicial subdivision $\mathcal{G}$ of $\mathcal{C}^n$ has to be such that it provides for a simplicial subdivision of each nonempty subset $A(s)$ of $\mathcal{C}^n$. Let t be the dimension of $A(s)$. Then $t = \mid \mathcal{I}^0(s) \mid +1$ where $\mathcal{I}^0(s) := \{i \in \mathcal{I}_n \mid s_i = 0\}$. So, if $x \in A(s)$ then there are a t-simplex $\sigma(y^1, \ldots, y^{t+1})$ in $A(s)$ and numbers $\lambda_1, \ldots, \lambda_{t+1} \geq 0$ such that $x = \sum_{i=1}^{t+1} \lambda_i y^i$ and $\sum_{i=1}^{t+1} \lambda_i = 1$, i.e. x is contained in $\sigma(y^1, \ldots, y^{t+1})$.

On the other hand, if $\operatorname{sgn}(F(x)) = s$, then there exist $\mu_h \geq 0$, $h \notin \mathcal{I}^0(s)$, such that $F(x) = \sum_{h \notin \mathcal{I}^0(s)} \mu_h s_h e(h)$. Hence, if x lies on the path $\overline{\mathcal{P}}$, then for a sign vector $s \in \mathbf{R}^n$ such that $s = \operatorname{sgn}(F(x))$ there exists a t-simplex $\sigma(y^1, \ldots, y^{t+1})$ in $A(s)$ such that the system of linear equations given by

$$\sum_{i=1}^{t+1} \lambda_i \begin{pmatrix} f(y^i) \\ 1 \end{pmatrix} - \sum_{h \notin \mathcal{I}^0(s)} \mu_h s_h \begin{pmatrix} e(h) \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{6.4}$$

has a nonnegative solution $\lambda_i \geq 0$, $i \in \mathcal{I}_{t+1}$, $\mu_h \geq 0$, $h \notin \mathcal{I}^0(s)$, with $x = \sum_{i=1}^{t+1} \lambda_i y^i$.

System (6.4) is a system of $n+1$ equations with $n+2$ unknowns leaving us with one degree of freedom. So, assuming nondegeneracy, the set of solutions if nonempty is a line segment which can be followed by making a linear programming pivot step in (6.4) with the column vector corresponding to the variable being zero in one of its end points. This line segment corresponds to a linear piece of $\overline{\mathcal{P}}$ in σ defined by the points $x = \sum_{i=1}^{t+1} \lambda_i y^i$. In the other end point, say $(\overline{\lambda}, \overline{\mu})$, of a nonempty line segment of solutions to (6.4) either $\overline{\lambda}_p = 0$ for some $p \in \mathcal{I}_{t+1}$ or $\overline{\mu}_k = 0$ for some $k \notin \mathcal{I}^0(s)$. Let $\overline{x} = \sum_{i=1}^{t+1} \overline{\lambda}_i y^i$.

Case 1: $\overline{\lambda}_p$ is 0 for some $p \in \mathcal{I}_{t+1}$. Then the end point $\overline{x} = \sum_{i \neq p} \overline{\lambda}_i y^i$ lies in the facet τ of σ opposite the vertex y^p. The facet τ is either also a facet of exactly one other t-simplex, say $\overline{\sigma}$, in $A(s)$ or τ lies in the boundary of $A(s)$.

Suppose $\overline{\sigma}$ exists. Then, in order to continue the path $\overline{\mathcal{P}}$ in $A(s)$, a pivot step is made in (6.4) with the column $(f(\overline{y})^\mathsf{T}, 1)^\mathsf{T}$ corresponding to the unique vertex $\overline{y}$ of $\overline{\sigma}$ not contained in τ.

Suppose $\overline{\sigma}$ does not exist and hence τ lies in the boundary of $A(s)$. If τ lies in $\mathcal{C}^n(s)$, then the algorithm has found a point $\overline{x} \in \mathcal{C}^n(s)$ with sign vector s equal to $\operatorname{sgn}(F(\overline{x}))$ so that $\overline{x}$ is an approximating solution for (6.1). Otherwise, τ is a $(t-1)$-simplex in $A(\overline{s})$ where $\overline{s}$ is a sign vector such that $\overline{s}_l \neq 0$ for some $l \in \mathcal{I}^0(s)$ while $\overline{s}_i = s_i$ for all $i \neq l$. Then the algorithm continues in $A(\overline{s})$ by pivoting the column $-\overline{s}_l(e(l)^\mathsf{T}, 0)^\mathsf{T}$ into (6.4).

Case 2: $\overline{\mu}_k$ is zero for some $k \notin \mathcal{I}^0(s)$. Then $F_k(\overline{x}) = s_k \overline{\mu}_k = 0$. Let $\overline{s}$ be a sign vector such that $\overline{s}_k = 0$ and $\overline{s}_h = s_h$ for $h \neq k$. Suppose that $A(\overline{s}) = \emptyset$. Then $\overline{x}$ lies in $\mathcal{C}^n(\overline{s})$ whereas $\operatorname{sgn}(F(\overline{x})) = \overline{s}$. Hence, $\overline{x}$ is an approximating solution to (6.1). Otherwise, if $A(\overline{s}) \neq \emptyset$, then there is exactly one $(t+1)$-simplex $\overline{\sigma}$ in $A(\overline{s})$ having σ as a facet. Now the algorithm continues by pivoting the column $(f(\overline{y})^\mathsf{T}, 1)^\mathsf{T}$ into

(6.4), where $\overline{y}$ is the vertex of $\overline{\sigma}$ not contained in σ.

Now we have described how the algorithm generates the path $\overline{\mathcal{P}}$ in the different subsets $A(s)$ of $\mathcal{C}^n$ we still have to describe the initialization of the algorithm at v. At v the system (6.4) equals

$$\lambda_1 \begin{pmatrix} f(v) \\ 1 \end{pmatrix} - \sum_{h=1}^{n} s_h^0 \mu_h \begin{pmatrix} e(h) \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{6.5}$$

having a unique solution $\lambda_1 = 1,\ \mu_h = s_h^0 f_h(v) > 0,\ h \in \mathcal{I}_n$, where $s^0 = \text{sgn}(f(v))$. If $A(s^0) = \emptyset$, then $v \in \mathcal{C}^n(s^0)$ and the algorithm stops with an exact solution to (6.1) at v. Otherwise, the starting point v is a facet of a unique 1-simplex $\sigma(y^1, y^2)$ in $A(s^0)$ with $y^1 = v$. The algorithm then pivots the column $(f(y^2)^\top, 1)^\top$ into (6.5).

Since all steps are unique, returning to v is impossible, and the number of simplices is finite, the algorithm terminates within a finite number of steps with an approximating solution $\overline{x}$ of (6.1). The accuracy of the approximation $f(\overline{x})$ can be measured by the smallest $\epsilon > 0$ for which for all $i \in \mathcal{I}_n$

$$\begin{aligned} f_i(\overline{x}) &\leq \epsilon \text{ if } a_i = \overline{x}_i \\ -\epsilon \leq f_i(\overline{x}) &\leq \epsilon \text{ if } a_i < \overline{x}_i < b_i \\ -\epsilon \leq f_i(\overline{x}) & \qquad\quad \overline{x}_i = b_i. \end{aligned}$$

If $f(\overline{x})$ is not accurate enough, i.e. if ϵ is too large, the algorithm is repeated being started at $v = \overline{x}$ with a finer simplicial subdivision of $\mathcal{C}^n$. This in the hope to find a more accurate approximation within a relative small number of steps. In this way, within a finite number of steps, an approximating solution with any accuracy can be found.

6.3 A simplicial subdivision

In order to subdivide $\mathcal{C}^n$ into simplices one can use any simplicial subdivision. The only restriction one has to pose on the simplicial subdivision of $\mathcal{C}^n$ to underly the algorithm described in Section 2 is that it provides for a simplicial subdivision of all nonempty subsets $A(s)$ of $\mathcal{C}^n$. A simplicial subdivision that perfectly fits into

this framework is the V-triangulation developed in Doup and Talman (1987) for the product space of unit simplices. In this section we adapt this V-triangulation to a simplicial subdivision of $\mathcal{C}^n$ meeting the requirements imposed on it.

To describe the simplicial subdivision of $\mathcal{C}^n$ we first subdivide each nonempty set $A(s)$ into subsets $A(s,\mathcal{T})$ where $\mathcal{T}$ is a subset of $\mathcal{I}^0(s) \cup -\mathcal{I}^0(s)$ such that for all $j \in I^0(s)$ either j or $-j$ belongs to $\mathcal{T}$.

Definition 6.3.1 *For each sign vector $s \in \mathbf{R}^n$ and each subset $\mathcal{T}$ of $\mathcal{I}^0(s) \cup -\mathcal{I}^0(s)$ such that either i or $-i$ belongs to $\mathcal{T}$,*

$$A(s,\mathcal{T}) = \emptyset \text{ if } v \in \mathcal{C}^n(s) \text{ and } (v_i = a_i \text{ for some } -i \in \mathcal{T} \text{ or } v_i = b_i \text{ for some } i \in \mathcal{T})$$

and otherwise

$$\begin{aligned} A(s,\mathcal{T}) = \{x \in \mathbf{R}^n_+ \mid\ & \text{if } s_i = +1 \text{ then} && x_i &=& (1-\rho)v_i + \rho b_i, \\ & \text{if } s_i = 0 \text{ and } i \in \mathcal{T} \text{ then} & v_i \le & x_i &\le& (1-\rho)v_i + \rho b_i, \\ & \text{if } s_i = 0 \text{ and } -i \in \mathcal{T} \text{ then} & (1-\rho)v_i + \rho a_i \le & x_i &\le& v_i, \\ & \text{if } s_i = -1 \text{ then} & (1-\rho)v_i + \rho a_i &=& x_i, \\ & \text{with } 0 \le \rho \le 1\}. \end{aligned}$$

Figure 6.3.1 gives a subdivision of $\mathcal{C}^n$ into subsets $A(s,\mathcal{T})$ for sign vectors $s \in \mathbf{R}^n$ when $n = 2$.

For some positive integer m, each nonempty set $A(s,\mathcal{T})$ is then subdivided into t-simplices $\sigma(y^1,\pi)$ with vertices $y^1,\ldots,y^{t+1}$ in $\mathcal{C}^n$, where $t = \mid \mathcal{I}^0(s) \mid + 1$ is the dimension of $A(s,\mathcal{T})$, such that

i) $y^1 = v + d(0)m^{-1}q(0) + \sum_{j\in\mathcal{I}^0(s)} d(j)m^{-1}q(j)$ with integers $d(j)$ and $d(0)$ satisfying

$$0 \le d(j) \le d(0) \le m-1 \text{ for all } j \in \mathcal{I}^0(s);$$

ii) $\pi = (\pi_1,\ldots,\pi_t)$ is a permutation of the elements of $\mathcal{I}^0(s) \cup \{0\}$ such that for all $j \in \mathcal{I}^0(s)$ it holds that

$$p > p' \text{ if } \pi_{p'} = 0,\ \pi_p = j,\ \text{and } d(\pi_{p'}) = d(\pi_p);$$

iii) $y^{i+1} = y^i + m^{-1}q(\pi_i),\ i = 1,\ldots,t,$

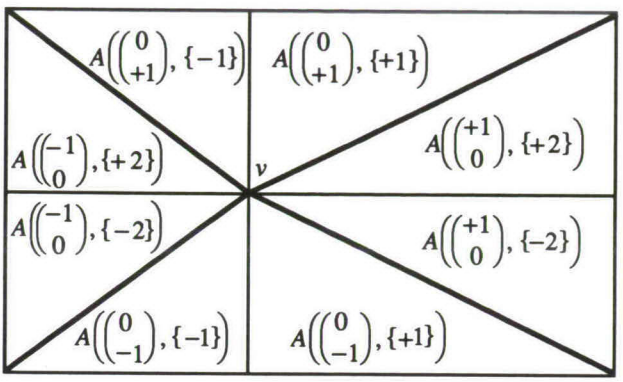


FIGURE 6.3.1: Subdivision of $\mathcal{C}^2$ into subsets $A(s, \mathcal{T})$.

where $q(0) = \sum_{j \in \mathcal{I}^{+1}(s) \cup \mathcal{T}} b_j e(j) + \sum_{j \in \mathcal{I}^{-1}(s) \cup -\mathcal{T}} a_j e(j) - v$ and

$$q(j) = (v_j - b_j)e(j) \text{ if } j \in \mathcal{T}$$
$$q(j) = (v_j - a_j)e(j) \text{ if } -j \in \mathcal{T}$$

for all $j \in \mathcal{I}^0(s)$.

If we denote this simplicial subdivision of $A(s, \mathcal{T})$ by $\mathcal{G}_m(s, \mathcal{T})$, then the set $A(s)$ is subdivided into simplices by the union $\mathcal{G}_m(s)$ of $\mathcal{G}_m(s, \mathcal{T})$ over all feasible $\mathcal{T}$. Moreover, the underlying V-triangulation of $\mathcal{C}^n$ is given by the union $\mathcal{G}_m$ of $\mathcal{G}_m(s)$ over all feasible s, m^{-1} being the grid size. Figure 6.3.2 gives a simplicial subdivision of $\mathcal{C}^n$ when $n = 2$ and $m = 2$.

In Section 2 we described how to follow the path $\overline{\mathcal{P}}$ through $\mathcal{C}^n$ by making pivot steps in the system of equations (6.4) with respect to a sequence of adjacent simplices in $A(s)$ for varying sign vectors s. After having introduced a specific simplicial subdivision of $\mathcal{C}^n$ we now describe how, given the parameters y^1, π, $d(h)$ for $h \in \mathcal{I}^0(s) \cup \{0\}$ of a t-simplex σ in $A(s)$ the parameters of a simplex $\overline{\sigma}$ in $A(s)$ adjacent to σ are obtained.

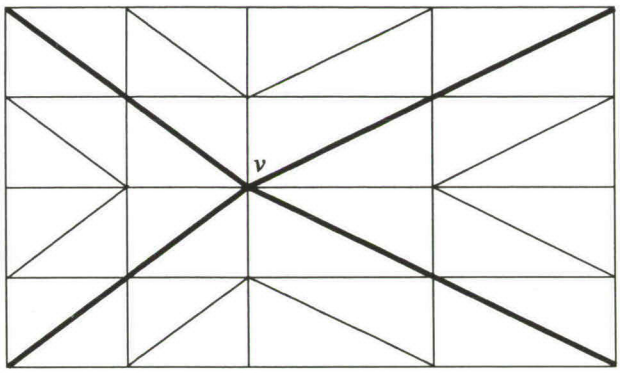


FIGURE 6.3.2: Simplicial subdivision of $\mathcal{C}^2$ with grid size $\frac{1}{2}$.

The movement from a t-simplex $\sigma(y^1, \pi)$ in $A(s, \mathcal{T})$ to an adjacent simplex $\overline{\sigma}(\overline{y}^1, \overline{\pi})$ is called a replacement step when $\overline{\sigma}(\overline{y}^1, \overline{\pi})$ is also a t-simplex in $A(s, \mathcal{T})$. Making a replacement step we replace the vertex y^p, for some $p \in \mathcal{I}_{t+1}$, of σ opposite the common facet τ of σ and $\overline{\sigma}$ by the vertex $\overline{y}$ of $\overline{\sigma}$ not belonging to τ. The possibilities are listed in Table 6.3.1, where the $(n+1)$-vector $d = (d_0, d_1, \ldots, d_n)^\mathsf{T}$ is defined by $d_h = d(h)$ for $h \in \mathcal{I}^0(s) \cup \{0\}$ and $d_h = 0$ otherwise. The $(n+1)$-dimensional unit vector $e(h)$ in Table 6.3.1 is given by $e_h(h) = 1$ and $e_i(h) = 0$ for $i \neq h$, $h = 0, 1, \ldots, n$.

TABLE 6.3.1: Replacement step

	$\overline{y}^1$	$\overline{\pi}$	$\overline{d}$
$p = 1$	$y^1 + m^{-1}q(\pi_1)$	$(\pi_2, \ldots, \pi_t, \pi_1)$	$d + e(\pi_1)$
$1 < p < t+1$	y^1	$(\pi_1, \ldots, \pi_{p-2}, \pi_p, \pi_{p-1}, \pi_{p+1}, \ldots, \pi_t)$	d
$p = t+1$	$y^1 - m^{-1}q(\pi_t)$	$(\pi_t, \pi_1, \ldots, \pi_{t-1})$	$d - e(\pi_t)$

In case the replacement step with respect to y^p cannot be performed, the facet τ of $\sigma(y^1, \pi)$ opposite y^p lies in the boundary of $A(s, \mathcal{T})$. Lemma 6.3.1 describes the cases when τ lies in the boundary of $A(s, \mathcal{T})$.

Lemma 6.3.1 *Let $\sigma(y^1, \pi)$ be a t-simplex in $\mathcal{G}_m(s, \mathcal{T})$ and let τ be the facet of σ opposite vertex y^p, $1 \leq p \leq t+1$. Then τ lies in the boundary of $A(s, \mathcal{T})$ if and only if one of the following cases holds:*

1) $p = 1$, $\pi_1 = 1$, and $d(\pi_1) = m - 1$;

2) $1 < p < t+1$, $\pi_{p-1} = 0$ and $\pi_p = j$ for some $j \in \mathcal{I}^0(s)$, and $d(\pi_{p-1}) = d(\pi_p)$;

3) $p = t+1$ and $d(\pi_t) = 0$.

When τ lies in the boundary of $A(s, \mathcal{T})$ then either τ lies in $\mathcal{C}^n(s)$ or in $A(\overline{s})$ for some $\overline{s}$ with $\overline{s}_j \neq 0$ for some $j \in \mathcal{I}^0(s)$ and $\overline{s}_h = s_h$ for $h \neq j$, or τ is a facet of an adjacent simplex in some set $A(s, \overline{\mathcal{T}})$ adjacent to $A(s, \mathcal{T})$, as follows.

Case 1: τ lies in $\mathcal{C}^n(s)$.

Case 2: If $v_j = a_j$ and $j \in \mathcal{T}$ ($v_j = b_j$ and $-j \in \mathcal{T}$) then τ is a $(t-1)$-simplex $\overline{\sigma}(y^1, \overline{\pi})$ in $A(\overline{s}, \overline{\mathcal{T}})$ where $\overline{\mathcal{T}} = \mathcal{T} \setminus \{j\}$ ($\overline{\mathcal{T}} = \mathcal{T} \setminus \{-j\}$), $\overline{s} = s - e(j)$ ($\overline{s} = s + e(j)$), and $\overline{\pi} = (\pi_1, \ldots, \pi_{p-1}, \pi_{p+1}, \ldots \pi_t)$. Notice that $q(j)$ disappears and $q(0)$ becomes $q(0) + q(j)$. Otherwise $\sigma(y^1, \pi)$ shares τ with an adjacent t-simplex $\overline{\sigma}(y^1, \pi)$ in $A(s, \overline{\mathcal{T}})$ where $\overline{\mathcal{T}} = \mathcal{T} \setminus \{j\} \cup \{-j\}$ if $j \in \mathcal{T}$ ($\overline{\mathcal{T}} = \mathcal{T} \setminus \{-j\} \cup \{j\}$ if $-j \in \mathcal{T}$).

Case 3: When $\mathcal{I}^0(s) \neq \emptyset$ then the facet τ opposite the vertex y^{t+1} of σ is the $(t-1)$-simplex $\overline{\sigma}(y^1, \overline{\pi})$ in $A(\overline{s}, \overline{\mathcal{T}})$, where $\overline{s} = s + e(\pi_t)$ and $\overline{\mathcal{T}} = \mathcal{T} \setminus \{\pi_t\}$ if $\pi_t \in \mathcal{T}$, $\overline{s} = s - e(\pi_t)$ and $\overline{\mathcal{T}} = \mathcal{T} \setminus \{-\pi_t\}$ if $-\pi_t \in \mathcal{T}$, and $\overline{\pi} = (\pi_1, \ldots, \pi_{t-1})$. Otherwise we have that $t = 1$ and $d(0) = 0$ which means that $\tau = \{v\}$.

Finally, a t-simplex $\sigma(y^1, \pi)$ in $A(s, \mathcal{T})$ is a facet of exactly one $(t+1)$-simplex $\overline{\sigma}$ in a nonempty set $A(\overline{s})$ where $\overline{s}_k = 0$ for some $k \notin I^0(s)$ and $\overline{s}_i = s_i$ for all other $i \in \mathcal{I}_n$. More precisely, $\overline{\sigma}$ is equal to $\overline{\sigma}(y^1, \overline{\pi})$ and lies in $A(\overline{s}, \overline{\mathcal{T}})$. If $s_k = +1$ and $v_k \neq b_k$ then $\overline{\mathcal{T}} = \mathcal{T} \cup \{+k\}$ and $\overline{\pi} = (\pi_1, \ldots, \pi_t, k)$. If $s_k = +1$ and $v_k = b_k$ then $\overline{\mathcal{T}} = \mathcal{T} \cup \{-k\}$, $d(k) = d(0)$, and $\overline{\pi} = (\pi_1, \ldots, \pi_{p-1}, k, \pi_p, \ldots, \pi_t)$ where p is such that $\pi_p = 0$. If $s_k = -1$ and $v_k \neq a_k$ then $\overline{\mathcal{T}} = \mathcal{T} \cup \{-k\}$ and $\overline{\pi} = (\pi_1, \ldots, \pi_t, k)$. If $s_k = -1$

and $v_k = a_k$ then $\overline{T} = T \cup \{+k\}$, $d(k) = d(0)$, and $\overline{\pi} = (\pi_1, \ldots, \pi_p, k, \pi_{p+1}, \ldots, \pi_t)$ where p is such that $\pi_p = 0$. This completes the description of the steps of the algorithm when the V-triangulation with grid size m^{-1} is the underlying subdivision of C^n.

Chapter 7

The nonlinear complementarity problem

The *nonlinear complementarity problem* is frequently met when solving systems of nonlinear equations or computing equilibria and fixed points. It generalizes the linear complementarity problem presented in Chapter 4 by replacing the linear relation between the variables in (4.1) by a nonlinear one. So the nonlinear complementarity problem can be defined as follows:

> Given a continuous function f from $\mathbf{R}^n$ to $\mathbf{R}^n$, find an $x^* \in \mathbf{R}^n$ such that
>
> $$x^* \geq 0, \quad f(x^*) \geq 0, \quad x^{*\mathsf{T}} f(x^*) = 0. \tag{7.1}$$

In this chapter we propose a simplicial restart algorithm that subdivides the set on which the problem is defined into simplices and generates from an arbitrarily chosen starting point a piecewise linear path either leading to an approximating solution or diverging to infinity. Given the possible occurence of divergence we will give a convergence condition under which the algorithm will find an approximating solution. If the accuracy of the approximating solution is not sufficient the algorithm can be restarted at the approximating solution with a finer simplicial subdivision. The piecewise linear path generated by the algorithm is followed by a sequence of adjacent simplices of varying dimension. The algorithm introduced in this chapter is

a generalization of the algorithm introduced in the previous chapter.

This chapter is based on Kremers and Talman (1990b) and is subdivided as follows. In Section 1 we describe the path to be followed by the algorithm while Section 2 describes the performance of the algorithm in the end points of the piecewise linear path. Also a convergence condition will be given in Section 2 guaranteeing the existence of an upper bound to the points generated by the algorithm. Section 3 describes an example of an appropriate simplicial subdivision of $\mathbf{R}^n_+$ to underly the algorithm.

7.1 The path to be approximated by the algorithm

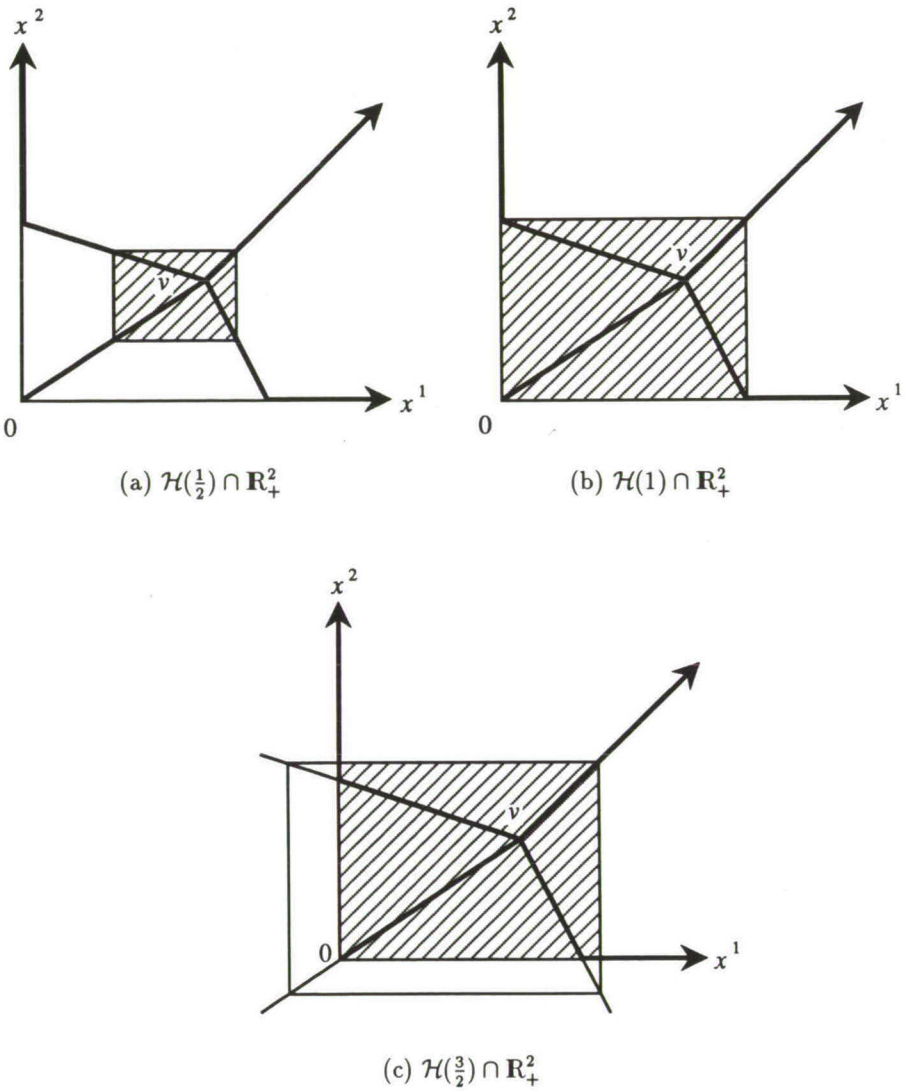
Starting in an arbitrarily chosen point $v \in \mathbf{R}^n_+$ the algorithm follows approximately a path in $\mathbf{R}^n_+$ such that each point x on the path is a stationary point of $-f$ on the parametrized set $\mathcal{H}(\rho) \cap \mathbf{R}^n_+$ for some ρ, $\rho \geq 0$. This set $\mathcal{H}(\rho)$, $\rho \geq 0$, is defined as

$$\mathcal{H}(\rho) = \{x \mid (1-\rho)v \leq x \leq v + \rho e\}. \tag{7.2}$$

For $\rho = 0$ the set $\mathcal{H}(0) \cap \mathbf{R}^n_+$ only consists of the starting point v. Hence v is a trivial solution to the stationary point problem of $-f$ on $\mathcal{H}(0) \cap \mathbf{R}^n_+$. For $0 < \rho < 1$, the set $\mathcal{H}(\rho) \cap \mathbf{R}^n_+$ is the cube $\{x \mid (1-\rho)v \leq x \leq v + \rho e\}$, and for $\rho = 1$ the set $\mathcal{H}(1) \cap \mathbf{R}^n_+$ is the cube $\{x \mid 0 \leq x \leq v + e\}$. If the algorithm generates a stationary point x of $-f$ on $\mathcal{H}(1) \cap \mathrm{bd}\mathbf{R}^n_+$ such that $x_i < v_i + 1$ for all $i \in \mathcal{I}_n$ then x is also a stationary point of $-f$ on $\mathbf{R}^n_+$ and thereby a solution to the nonlinear complementarity problem. For $\rho \geq 1$ the set $\mathcal{H}(\rho) \cap \mathbf{R}^n_+$ is equal to the cube $\{x \mid 0 \leq x \leq v + \rho e\}$. Clearly the set $\mathcal{H}(\rho) \cap \mathbf{R}^n_+$ is equal to the set

$$\{x \in \mathbf{R}^n \mid \max\{0, (1-\rho)v_i\} \leq x_i \leq v_i + \rho \text{ for } i \in \mathcal{I}_n\}.$$

For an illustration of the set $\mathcal{H}(\rho) \cap \mathbf{R}^n_+$ we refer to Figure 7.1.1. The equivalence between the stationary point problem on this set and the nonlinear complementarity problem with lower bounds $\max\{0, (1-\rho)v_i\}$ and upper bounds $v_i + \rho$ for $i \in \mathcal{I}_n$ as proved in Theorem 6.1.1 of the previous chapter implies that at every point x on the



(a) $\mathcal{H}(\frac{1}{2}) \cap \mathbf{R}^2_+$

(b) $\mathcal{H}(1) \cap \mathbf{R}^2_+$

(c) $\mathcal{H}(\frac{3}{2}) \cap \mathbf{R}^2_+$

FIGURE 7.1.1: The subset $\mathcal{H}(\rho) \cap \mathbf{R}^2_+$ for $\rho = \frac{1}{2},\ 1,\ \frac{3}{2}$, given $v = (3,2)^\top$.

path for all $i \in \mathcal{I}_n$ it holds that

$$\begin{aligned} f_i(x) &\geq 0 \text{ if } \max\{0,(1-\rho)v_i\} = x_i \\ f_i(x) &= 0 \text{ if } \max\{0,(1-\rho)v_i\} < x_i < v_i + \rho \\ f_i(x) &\leq 0 \text{ if } x_i = v_i + \rho \end{aligned} \tag{7.3}$$

for some $\rho \geq 0$. Under certain regularity and nondegeneracy conditions, see van den Elzen (1991), the set of points x satisfying the conditions in (7.3) for $\rho \geq 0$ consists of piecewise smooth curves. Each of these curves is either a loop or a path. One of these paths, say $\mathcal{P}$, has v as an end point for $\rho = 0$. If $\mathcal{P}$ has another end point, say x^*, then x^* is a solution to (7.1). Otherwise the path $\mathcal{P}$ will go to infinity.

Without loss of generality we assume that no component of $f(v)$ equals zero. Then, by increasing ρ from zero the path $\mathcal{P}$ leaves v through increasing x_i from v_i such that $x_i = v_i + \rho$ if $f_i(v) < 0$ and through decreasing x_i from v_i such that $x_i = \max\{0,(1-\rho)v_i\}$ if $f_i(v) > 0$, for all $i \in \mathcal{I}_n$. If along the path $\mathcal{P}$, at a point x satisfying (7.3), $f_j(x)$ becomes zero for some $j \in \mathcal{I}_n$ while $x_j = v_j + \rho$ (or $x_j = \max\{0,(1-\rho)v_j\}$) then either x solves (7.1) or the path continues by decreasing (increasing) x_j from $v_j + \rho$ ($\max\{0,(1-\rho)v_j\}$) and keeping $f_j(x) = 0$. If at a point x on $\mathcal{P}$, satisfying (7.3), x_j becomes equal to $v_j + \rho$ (or $\max\{0,(1-\rho)v_j\}$), for some $j \in \{i \mid f_i(x) = 0\}$, then the path $\mathcal{P}$ continues by decreasing (increasing) $f_j(x)$ from 0 and keeping x_j equal to $v_j + \rho$ (or $\max\{0,(1-\rho)v_j\}$). Finally, if at a point x on $\mathcal{P}$, ρ becomes equal to 1 and $f(x) \geq 0$, then $f_i(x) = 0$ when $0 < x_i \leq v_i + 1$ and $f_i(x) \geq 0$ when $x_i = 0$, so x is a solution of (7.1). Otherwise, ρ will be increased further, keeping $x_i = 0$ for all i such that $f_i(x) > 0$. In the next section we give an algorithm to follow the path $\mathcal{P}$ approximately.

7.2 The algorithm

The algorithm approximately follows the path $\mathcal{P}$ described in Section 1 by generating a piecewise linear path $\overline{\mathcal{P}}$ connecting v with an approximating solution $\overline{x}$ of (7.1) or diverging towards infinity. For a description of this piecewise linear path we approximate the function f by a piecewise linear approximation F of f on a simplicial subdivision of $\mathbf{R}^n_+$, which we denote by $\mathcal{G}$. For an appropriate simplicial subdivision of $\mathbf{R}^n_+$ we refer to Section 3 of this chapter.

The results obtained in Section 1 of this chapter with respect to f can also be applied to the piecewise linear approximation F of f on $\mathcal{G}$. In particular, there exists a piecewise linear path $\overline{\mathcal{P}}$ of points in $\mathbf{R}^n_+$ starting in v and ending in a solution to (7.1) with respect to F or going to infinity. For each point x on the path $\overline{\mathcal{P}}$ there exists a $\rho \geq 0$ such that for all $i \in \mathcal{I}_n$

$$\begin{aligned} \max\{0,(1-\rho)v_i\} &= x_i && \text{if } F_i(x) > 0 \\ \max\{0,(1-\rho)v_i\} \leq x_i &\leq v_i + \rho && \text{if } F_i(x) = 0 \\ x_i &= v_i + \rho && \text{if } F_i(x) < 0. \end{aligned} \tag{7.4}$$

For each sign vector s let the sets $A(s)$ and $A^0(s)$ be defined as follows.

Definition 7.2.1 *For each sign vector* $s \in \mathbf{R}^n$

$$A(s) = \emptyset \text{ if } s \geq 0 \text{ and } v_i = 0 \text{ for all } i \in \mathcal{I}^{+1}(s),$$

and otherwise

$$\begin{aligned} A(s) = \{x \in \mathbf{R}^n_+ \mid\ & \text{if } s_i = +1 \text{ then } \max\{0,(1-\rho)v_i\} = x_i, \\ & \text{if } s_i = 0 \text{ then } \max\{0,(1-\rho)v_i\} \leq x_i \leq v_i + \rho, \\ & \text{if } s_i = -1 \text{ then } x_i = v_i + \rho, \\ & \text{with } \rho \geq 0 \text{ if } s \leq 0 \text{ or } v_i = 0 \text{ for all } i \in \mathcal{I}^{+1}(s), \\ & \text{and otherwise } 0 \leq \rho \leq 1\}. \end{aligned}$$

Definition 7.2.2 *For each sign vector* $s \in \mathbf{R}^n$

$$A^0(s) = \emptyset \text{ if } s \leq 0 \text{ or } s \geq 0 \text{ or } v_i = 0 \text{ for all } i \in \mathcal{I}^{+1}(s),$$

and otherwise

$$\begin{aligned} A^0(s) = \{x \in \mathbf{R}^n_+ \mid\ & \text{if } s_i = +1 \text{ then } 0 = x_i, \\ & \text{if } s_i = 0 \text{ then } 0 \leq x_i \leq v_i + \rho, \\ & \text{if } s_i = -1 \text{ then } x_i = v_i + \rho, \text{ with } \rho \geq 1\}. \end{aligned}$$

Figure 7.2.1 gives a subdivision of $\mathbf{R}^n_+$ into subsets $A(s)$ and $A^0(s)$ for sign vectors $s \in \mathbf{R}^n$ when $n = 2$. Each point x on the path $\overline{\mathcal{P}}$ generated by the algorithm is then for some sign vector $s \in \mathbf{R}^n$ a point in $A(s)$ or in $A^0(s)$ such that $s = \operatorname{sgn}(F(x))$.

The simplicial subdivision $\mathcal{G}$ of $\mathbf{R}^n_+$ must be such that it gives a simplicial subdivision of each nonempty subset $A(s)$ and $A^0(s)$ of $\mathbf{R}^n_+$. Let t be the dimension of

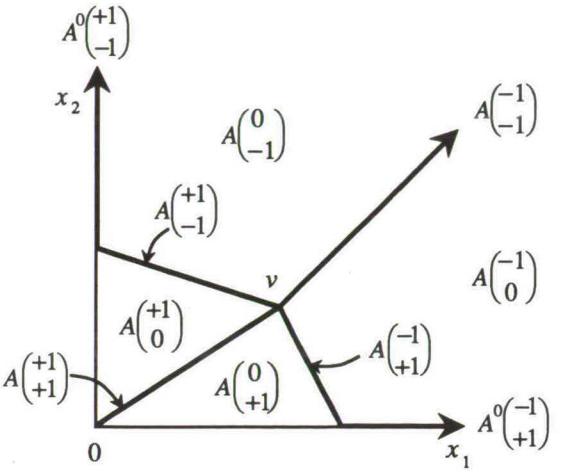


FIGURE 7.2.1: Subdivision of $\mathbf{R}^2_+$ into subsets $A(s)$ and $A^0(s)$ for sign vectors $s \in \mathbf{R}^2$.

a nonempty $A(s)$ (or $A^0(s)$). Then $t = \mid \mathcal{I}^0(s) \mid +1$. So, if $x \in A(s)$ (or $x \in A^0(s)$), there are a t-simplex $\sigma(y^1, \ldots, y^{t+1})$ in $A(s)$ ($A^0(s)$) and numbers $\lambda_1, \ldots, \lambda_{t+1} \geq 0$ such that $x = \sum_{i=1}^{t+1} \lambda_i y^i$ and $\sum_{i=1}^{t+1} \lambda_i = 1$, i.e. $\sigma(y^1, \ldots, y^{t+1})$ contains x. On the other hand, if $\text{sgn}(F(x)) = s$, then there exist $\mu_h \geq 0$, $h \notin \mathcal{I}^0(s)$, such that $F(x) = \sum_{h \notin \mathcal{I}^0(s)} \mu_h s_h e(h)$. Hence if $x \in A(s)$ ($x \in A^0(s)$) for some sign vector $s \in \mathbf{R}^n$ such that $s = \text{sgn}(F(x))$ then there exists a t-simplex $\sigma(y^1, \ldots, y^{t+1})$ in $A(s)$ ($A^0(s)$) such that the system of linear equations given by

$$\sum_{i=1}^{t+1} \lambda_i \begin{pmatrix} f(y^i) \\ 1 \end{pmatrix} - \sum_{h \notin \mathcal{I}^0(s)} \mu_h s_h \begin{pmatrix} e(h) \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{7.5}$$

has a solution $\lambda_i \geq 0$, $i \in \mathcal{I}_{t+1}$, $\mu_h \geq 0$, $h \notin \mathcal{I}^0(s)$, such that $x = \sum_{i=1}^{t+1} \lambda_i y^i$.

System (7.5) is a system of $n+1$ equations with $n+2$ unknowns leaving us with one degree of freedom. So, assuming nondegeneracy, the set of solutions to (7.5) if it exists is a line segment. It can be followed by making a linear programming pivot step in (7.5) with the column corresponding to the variable being zero in one of its end points. This line segment corresponds to a linear piece of $\overline{\mathcal{P}}$ in σ defined by the points $x = \sum_{i=1}^{t+1} \lambda_i y^i$.

In an end point $(\overline{\lambda}, \overline{\mu})$ of a line segment of solutions to (7.5) either $\overline{\lambda}_p = 0$ for some $p \in \mathcal{I}_{t+1}$ or $\overline{\mu}_j = 0$ for some $j \notin \mathcal{I}^0(s)$. Let $\overline{x} = \sum_{i=1}^{t+1} \overline{\lambda}_i y^i$.

Case 1: $\overline{\lambda}_p$ is equal to 0 for some $p \in \mathcal{I}_{t+1}$. Then $\overline{x} = \sum_{i \neq p} \overline{\lambda}_i y^i$ lies in the facet τ of σ opposite the vertex y^p. The facet τ either is a facet of exactly one other t-simplex, say $\overline{\sigma}$, in $A(s)$ $(A^0(s))$, or lies in the boundary of $A(s)$ $(A^0(s))$.

Suppose $\overline{\sigma}$ exists. Then, in order to continue the path $\overline{\mathcal{P}}$ in $A(s)$ $(A^0(s))$, a pivot step is made in (7.5) with the column $(f(\overline{y})^\mathsf{T}, 1)^\mathsf{T}$ corresponding to the unique vertex $\overline{y}$ of $\overline{\sigma}$ not contained in τ. The algorithm is continued by repeating the described procedure.

Suppose $\overline{\sigma}$ does not exist, and hence τ lies in the boundary of $A(s)$ $(A^0(s))$. Let the set $\mathbf{R}^n(s)$ in the boundary of $A(s)$ or $A^0(s)$ be defined as follows.

Definition 7.2.3 *For each sign vector* $s \in \mathbf{R}^n$ *let*

$$\mathbf{R}^n(s) = \emptyset \text{ if } s \leq 0 \text{ or } v_i = 0 \text{ for all } i \in \mathcal{I}^{+1}(s),$$

and otherwise

$$\begin{aligned} \mathbf{R}^n(s) = \{x \in \mathbf{R}^n_+ \mid\ & \text{if } s_i = +1 \quad \text{then } 0 = x_i, \\ & \text{if } s_i = 0 \quad \text{then } 0 \leq x_i \leq v_i + 1, \\ & \text{if } s_i = -1 \quad \text{then} \quad x_i = v_i + 1\}. \end{aligned}$$

If τ lies in $\mathbf{R}^n(s)$, then the algorithm has found a point $\overline{x} \in \mathbf{R}^n(s)$ with sign vector s equal to $\operatorname{sgn}(F(\overline{x}))$. In case $s \geq 0$ then $\overline{x}$ is an approximating solution to (7.1). Otherwise $\mathbf{R}^n(s) = A(s) \cap A^0(s)$ and τ is the facet of a unique t-simplex $\hat{\sigma}$ in $A^0(s)$ $(A(s))$. In order to continue the path $\overline{\mathcal{P}}$ in $A^0(s)$ $(A(s))$, a linear programming pivot step is made in (7.5) with the column vector $(f(\hat{y})^\mathsf{T}, 1)^\mathsf{T}$ corresponding to the unique vertex $\hat{y}$ of $\hat{\sigma}$ not contained in τ. The algorithm is continued by repeating the described procedure.

If the facet τ of σ in the boundary of $A(s)$ $(A^0(s))$ does not lie in $\mathbf{R}^n(s)$ then τ is a $(t-1)$-simplex in $A(\overline{s})$ or $A^0(\overline{s})$, with $\overline{s}$ being a sign vector such that $\overline{s}_l \neq 0$ for some unique $l \in \mathcal{I}^0(s)$ while $\overline{s}_i = s_i$ for all $i \neq l$, and the algorithm continues in $A(\overline{s})$ or $A^0(\overline{s})$ by pivoting the column vector $-\overline{s}_l(e(l)^\mathsf{T}, 0)^\mathsf{T}$ into (7.5).

Case 2: $\overline{\mu}_k$ is equal to 0 for some $k \notin \mathcal{I}^0(s)$. Hence, $F_k(\overline{x}) = s_k \overline{\mu}_k = 0$. Let $\overline{s}$ be a sign vector such that $\overline{s}_k = 0$ and $\overline{s}_h = s_h$ for $h \neq k$. Then $\overline{x}$ is an approximating

solution to (7.1) if $\overline{x} \in A(s)$, $\overline{s} \geq 0$ and $v_i = 0$ for all $i \in \mathcal{I}^{+1}(\overline{s})$. Also $\overline{x}$ is an approximating solution to (7.1) if $\overline{x} \in A^0(s)$ and $\overline{s} \geq 0$. Otherwise there is exactly one $(t+1)$-simplex $\overline{\sigma}$ in $A(\overline{s})$ or $A^0(\overline{s})$ having σ as a facet. The algorithm proceeds by pivoting the column $(f(\overline{y})^\top, 1)^\top$ into (7.5), where $\overline{y}$ is the vertex of $\overline{\sigma}$ not contained in σ.

Now we have described how the algorithm generates the path $\overline{\mathcal{P}}$ in the different subsets $A(s)$ and $A^0(s)$ of $\mathbf{R}^n_+$, we still have to describe the initialization of the algorithm at v. At v the system (7.5) becomes

$$\lambda_1 \begin{pmatrix} f(v) \\ 1 \end{pmatrix} - \sum_{h=1}^{n} s^0_h \mu_h \begin{pmatrix} e(h) \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{7.6}$$

having a unique solution $\lambda_1 = 1$, $\mu_h = s^0_h f_h(v) > 0$, $h \in \mathcal{I}_n$, where $s^0 = \operatorname{sgn}(f(v))$. If $s^0 \geq 0$ and $v_i = 0$ for all $i \in \mathcal{I}^{+1}(s^0)$, then the algorithm stops with an exact solution to (7.1) at v. Otherwise, the starting point v is a facet of a unique 1-simplex $\sigma(y^1, y^2)$ in $A(s^0)$ with $y^1 = v$. The algorithm then pivots the column $(f(y^2)^\top, 1)^\top$ into (7.6).

Since all steps are unique and returning to v is impossible the algorithm either terminates within a finite number of steps with an approximating solution $\overline{x}$ to (7.1) or it follows a path to infinity.

When an approximating solution $\overline{x}$ to (7.1) is found one can measure the accuracy of approximation by taking the smallest $\epsilon > 0$ for which for all $i \in \mathcal{I}_n$

$$\begin{aligned} -\epsilon \leq f_i(\overline{x}) \qquad & \text{if } \overline{x}_i = 0 \\ -\epsilon \leq f_i(\overline{x}) \leq \epsilon \quad & \text{if } \overline{x}_i > 0. \end{aligned}$$

If $f(\overline{x})$ is not accurate enough, i.e., ϵ is too large, the algorithm is repeated being started at $v = \overline{x}$ with a finer simplicial subdivision of $\mathbf{R}^n_+$. This in the hope to find a more accurate approximation within a few number of steps.

When a path towards infinity is followed no solution will be found. So, we have to state conditions on f under which the algorithm converges towards an approximating solution to (7.1). Theorem 7.2.1 states conditions under which an upper bound to the points generated by the algorithm exists.

Theorem 7.2.1 *Let f be a continuous function from $\mathbf{R}^n$ to $\mathbf{R}^n$. Assume there exists a number $\mu > 0$ such that $f_i(x) > 0$ whenever $x_i > \mu$, for all $i \in \mathcal{I}_n$. Then the algorithm terminates within a finite number of steps with a solution to (7.1) with respect to F.*

Proof: Without loss of generality we may assume that $1 + v_i < \mu$ for all $i \in \mathcal{I}_n$. We show that the algorithm cannot generate a t-simplex σ, with vertices $y^1, \ldots, y^{t+1}$, in the simplicial subdivision $\mathcal{G}$ of $\mathbf{R}^n_+$ for which for all $x \in \sigma$ it holds that $x_i > \mu$ for some $i \in \mathcal{I}_n$. Suppose it does, then, since $x_i > v_i + 1$ for all $x \in \sigma$, σ lies in $A(s)$ or $A^0(s)$ where s is a sign vector such that $s_i \leq 0$. We conclude from (7.5) that the i-th component of the piecewise linear approximation F in $x = \sum_{j=1}^{t+1} \lambda_j y^j \in \sigma$ is less than or equal to zero, i.e. $F_i(x) \leq 0$. On the other hand, by assumption, $y \in \sigma$ implies $y_i > \mu$, so $f_i(y^j) > 0$ for all vertices y^j, $j \in \mathcal{I}_{t+1}$, of σ. Since x is a convex combination of these vertices, the piecewise linear approximation F in x, being the same convex combination of $f(y^i)$, $j \in \mathcal{I}_{t+1}$, must have a positive i-th component, i.e. $F_i(x) > 0$. This is in contradiction with $F_i(x) \leq 0$. So, σ cannot be generated by the algorithm. □

This convergency condition is similar to the convergency condition given in Merrill (1971) for simplicial homotopy algorithms on $\mathbf{R}^n$. Merrill's condition implies that there exists a cube, or ball, such that the algorithm will not generate points outside this compact set. The convergency condition stated above caused us to let the algorithm follow a path of stationary points of $-f$ on an expanding cube instead of the set introduced in Chapter 5 to solve the linear complementarity problem. If the number μ is known in advance then the algorithm of this chapter coincides with the one of the previous chapter when taking $a = 0$, $b = \mu e$, and $-f$ instead of f. Moreover, $\mathcal{H}(\rho)$ then becomes equal to the set $\{x \in \mathbf{R}^n \mid (1-\rho)v \leq x \leq (1-\rho)v + \rho\mu e\}$ for $0 \leq \rho \leq 1$.

7.3 A simplicial subdivision

In order to subdivide $\mathbf{R}^n_+$ into simplices one can use any simplicial subdivision. The only restriction one has to pose on the simplicial subdivision of $\mathbf{R}^n_+$ to underly the algorithm described in Section 2 of this chapter is that it provides for a simplicial

subdivision of all nonempty subsets $A(s)$ and $A^0(s)$ of $\mathbf{R}^n_+$. In this section we propose an appropriate simplicial subdivision of $\mathbf{R}^n_+$ which is based on a combination of the V-triangulation and the K'-triangulation as described in Section 2 of Chapter 3. To describe this simplicial subdivision of $\mathbf{R}^n_+$ we need to subdivide each nonempty set $A(s)$ into subsets $A(s,\mathcal{T})$ and each nonempty set $A^0(s)$ into subsets $A^0(s,\mathcal{T})$ where $\mathcal{T}$ is a subset of $\mathcal{I}^0(s) \cup -\mathcal{I}^0(s)$ such that for all $i \in \mathcal{I}^0(s)$ either i or $-i$ belongs to $\mathcal{T}$.

Definition 7.3.1 *For each sign vector $s \in \mathbf{R}^n$ such that $s \not\geq 0$ or $v_i > 0$ for some $i \in \mathcal{I}^{+1}(s)$, and for each subset $\mathcal{T}$ of $\mathcal{I}^0(s) \cup -\mathcal{I}^0(s)$ such that either i or $-i$ belongs to $\mathcal{T}$,*

$$A(s,\mathcal{T}) = \emptyset \text{ if } v_i = 0 \text{ for some } i \in \mathcal{T},$$

and otherwise

$$\begin{aligned} A(s,\mathcal{T}) = \{x \in \mathbf{R}^n_+ \mid\ & \text{if } s_i = +1 \text{ then} && \max\{0,(1-\rho)v_i\} &&= x_i, \\ & \text{if } s_i = 0 \text{ and } i \in \mathcal{T} \text{ then} && \max\{0,(1-\rho)v_i\} \leq x_i &\leq v_i, \\ & \text{if } s_i = 0 \text{ and } -i \in \mathcal{T} \text{ then} && v_i \leq x_i &\leq v_i + \rho, \\ & \text{if } s_i = -1 \text{ then} && x_i &= v_i + \rho, \\ & \text{with } \rho \geq 0 \text{ if } s \leq 0 \text{ or } v_i = 0 \text{ for all } i \in \mathcal{I}^{+1}(s), \\ & \text{and otherwise } 0 \leq \rho \leq 1\}. \end{aligned}$$

Definition 7.3.2 *For each sign vector $s \in \mathbf{R}^n$ such that $s \not\leq 0$, $s \not\geq 0$, and $v_i > 0$ for some $i \in \mathcal{I}^{+1}(s)$, and for each subset $\mathcal{T}$ of $\mathcal{I}^0(s) \cup -\mathcal{I}^0(s)$ such that either i or $-i$ belongs to $\mathcal{T}$,*

$$A^0(s,\mathcal{T}) = \emptyset \text{ if } v_i = 0 \text{ for some } i \in \mathcal{T},$$

and otherwise

$$\begin{aligned} A^0(s,\mathcal{T}) = \{x \in \mathbf{R}^n_+ \mid\ & \text{if } s_i = +1 \text{ then} && 0 &= x_i, \\ & \text{if } s_i = 0 \text{ and } i \in \mathcal{T} \text{ then} && 0 \leq x_i &\leq v_i, \\ & \text{if } s_i = 0 \text{ and } -i \in \mathcal{T} \text{ then} && v_i \leq x_i &\leq v_i + \rho, \\ & \text{if } s_i = -1 \text{ then} && x_i &= v_i + \rho, \\ & \text{with } \rho \geq 1\}. \end{aligned}$$

Figure 7.3.1 gives a subdivision of $\mathbf{R}^n_+$ into subsets $A(s,\mathcal{T})$ and $A^0(s,\mathcal{T})$ when $n = 2$.

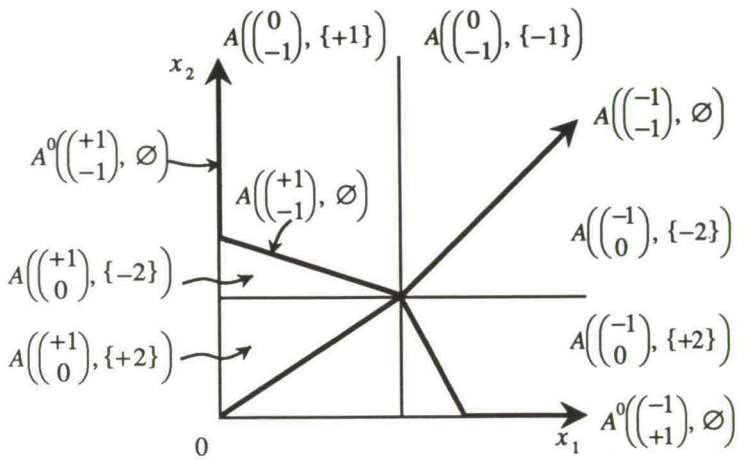


FIGURE 7.3.1: Subdivision of $\mathbf{R}^2_+$ into subsets $A(s,\mathcal{T})$ and $A^0(s,\mathcal{T})$.

For some positive integer m, each nonempty subset $A(s,\mathcal{T})$ is subdivided into t-simplices $\sigma(y^1,\pi)$ with vertices $y^1,\ldots,y^{t+1}$, where $t = |\,\mathcal{I}^0(s)\,| + 1$ is the dimension of $A(s,\mathcal{T})$, such that

i) $y^1 = v + d(0)m^{-1}q(0) + \sum_{j\in\mathcal{I}^0(s)} d(j)m^{-1}q(j)$ with integers $d(j)$ and $d(0)$ satisfying

if $s \leq 0$ or $v_i = 0$ for all $i \in \mathcal{I}^{+1}(s)$ then

$$0 \leq d(j) \leq d(0) \text{ for all } -j \in \mathcal{T}, \text{ and}$$

$$\max\{0, d(0) - m\} \leq d(j) \leq d(0) \text{ for all } j \in \mathcal{T},$$

and otherwise

$$0 \leq d(j) \leq d(0) \leq m - 1 \text{ for all } j \in \mathcal{I}^0(s);$$

ii) $\pi = (\pi_1, \ldots, \pi_t)$ is a permutation of the elements of $\mathcal{I}^0(s) \cup \{0\}$ such that for all $j \in \mathcal{I}^0(s)$:

$$\text{if } \pi_{p'} = 0,\ \pi_p = j, \text{ and } d(\pi_p) = d(\pi_{p'}) \text{ then } p' < p;$$

$$\text{if } \pi_{p'} = 0,\ \pi_p = j,\ j \in \mathcal{T},\ s \leq 0 \text{ or } v_i = 0 \text{ for all } i \in \mathcal{I}^{+1}(s),$$
$$\text{and } d(\pi_p) = d(\pi_{p'}) - m \text{ then } p < p';$$

iii) $y^{i+1} = y^i + m^{-1}q(\pi_i),\ \ i = 1, \ldots, t,$

where $q_j(0) = 1$ for $j \in \mathcal{I}^{-1}(s)$ or $-j \in \mathcal{T}$ and $q_j(0) = -v_j$ for $j \in \mathcal{I}^{+1}(s)$ or $j \in \mathcal{T}$, and where

$$q(j) = -e(j) \quad \text{if } -j \in \mathcal{T},$$
$$q(j) = v_j e(j) \quad \text{if } j \in \mathcal{T}.$$

Let the simplicial subdivision of $A(s, \mathcal{T})$ be denoted by $\mathcal{G}_m(s, \mathcal{T})$. Then, the simplicial subdivision of $A(s)$, denoted by $\mathcal{G}_m(s)$, is given by the union of $\mathcal{G}_m(s, \mathcal{T})$ over all feasible subsets $\mathcal{T}$.

For the same positive integer m, each nonempty subset $A^0(s, \mathcal{T})$ is subdivided into t-simplices $\sigma(y^1, \pi)$ with vertices $y^1, \ldots, y^{t+1}$, where $t = \mid \mathcal{I}^0(s) \mid + 1$ is the dimension of $A^0(s, \mathcal{T})$, such that

i) $y^1 = v(s, \mathcal{T}) + d(0)m^{-1}q(0) + \sum_{j \in \mathcal{I}^0(s)} d(j)m^{-1}q(j)$ with integers $d(j)$ and $d(0)$ satisfying

$$d(0) \geq m;$$

$$0 \leq d(j) \leq d(0) \text{ for all } -j \in \mathcal{T};$$

$$d(0) - m \leq d(j) \leq d(0) \text{ for all } j \in \mathcal{T};$$

ii) $\pi = (\pi_1, \ldots, \pi_t)$ is a permutation of the elements of $\mathcal{I}^0(s) \cup \{0\}$ such that for all $j \in \mathcal{I}^0(s)$:

$$\text{if } \pi_{p'} = 0,\ \pi_p = j, \text{ and } d(\pi_p) = d(\pi_{p'}) \text{ then } p' < p;$$

$$\text{if } \pi_{p'} = 0,\ \pi_p = j,\ j \in \mathcal{T}, \text{ and } d(\pi_p) = d(\pi_{p'}) - m \text{ then } p < p';$$

iii) $y^{i+1} = y^i + m^{-1}q(\pi_i), \quad i = 1, \ldots, t,$

where $v_j(s, \mathcal{T}) = 0$ for all $j \in \mathcal{I}^{+1}(s)$ and $v_j(s, \mathcal{T}) = v_j$ otherwise, where $q_j(0) = 1$ for $j \in \mathcal{I}^{-1}(s)$ or $-j \in \mathcal{T}$, $q_j(0) = -v_j$ for $j \in \mathcal{T}$, $q_j(0) = 0$ for $j \in \mathcal{I}^{+1}(s)$, and where

$$q(j) = -e(j) \quad \text{if } -j \in \mathcal{T},$$
$$q(j) = v_j e(j) \quad \text{if } j \in \mathcal{T}.$$

Let the simplicial subdivision of $A^0(s, \mathcal{T})$ be denoted by $\mathcal{G}_m^0(s, \mathcal{T})$. Then, the simplicial subdivision of $A^0(s)$, denoted by $\mathcal{G}_m^0(s)$, is given by the union of $\mathcal{G}_m^0(s, \mathcal{T})$ over all feasible $\mathcal{T}$. The simplicial subdivision of $\mathbf{R}_+^n$, $\mathcal{G}_m$, is now induced by the union of $\mathcal{G}_m(s)$ and $\mathcal{G}_m^0(s)$ over all possible sign vectors s, m^{-1} being the grid size of the simplicial subdivision. Figure 7.3.2 gives a simplicial subdivision of $\mathbf{R}_+^n$ for $n = 2$ and $m = 2$.

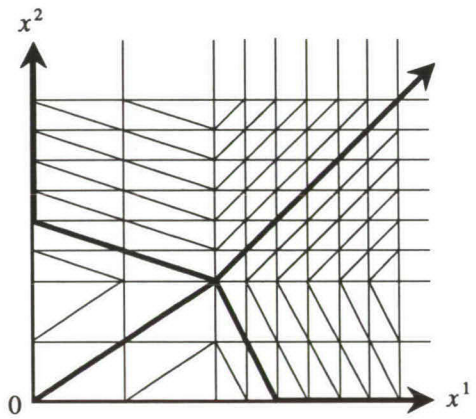


FIGURE 7.3.2: Simplicial subdivision of $\mathbf{R}_+^2$ with grid size $\frac{1}{2}$.

In Section 2 we described how to follow the path $\overline{\mathcal{P}}$ through $\mathbf{R}_+^n$ from v by making pivot steps in the system of equations (7.5) with respect to a sequence of adjacent simplices σ in $A(s)$ or $A^0(s)$ for varying sign vectors s. After having introduced a specific simplicial subdivision of $\mathbf{R}_+^n$ we now describe how a sequence of adjacent simplices σ in this simplicial subdivision of $\mathbf{R}_+^n$ can be followed, i.e., we describe

how, given the parameters y^1, π, and $d(h)$, for $h \in \mathcal{I}^0(s) \cup \{0\}$, of a t-simplex σ the parameters of a simplex $\overline{\sigma}$ adjacent to σ are obtained.

The movement from a t-simplex $\sigma(y^1, \pi)$ in $A(s, \mathcal{T})$ $(A^0(s, \mathcal{T}))$ to an adjacent simplex $\overline{\sigma}(\overline{y}^1, \overline{\pi})$ is called a replacement step when $\overline{\sigma}(\overline{y}^1, \overline{\pi})$ is also a t-simplex in $A(s, \mathcal{T})$ $(A^0(s, \mathcal{T}))$. Making a replacement step we replace a vertex y^p, for some $p \in \mathcal{I}_{t+1}$, of σ opposite the common facet τ of σ and $\overline{\sigma}$ by the vertex $\overline{y}$ of $\overline{\sigma}$ not belonging to τ. The possibilities are listed in Table 6.3.1.

In case the replacement step with respect to y^p cannot be performed the facet τ of $\sigma(y^1, \pi)$ opposite y^p lies in the boundary of $A(s, \mathcal{T})$ $(A^0(s, \mathcal{T}))$. Lemma 7.3.1 describes the cases when τ lies in the boundary of $A(s, \mathcal{T})$ or $A^0(s, \mathcal{T})$.

Lemma 7.3.1 *Let $\sigma(y^1, \pi)$ be a t-simplex in $\mathcal{G}_m(s, \mathcal{T})$ $(\mathcal{G}^0_m(s, \mathcal{T}))$ and let τ be the facet of σ opposite vertex y^p, $1 \leq p \leq t+1$. Then τ lies in the boundary of $A(s, \mathcal{T})$ $(A^0(s, \mathcal{T}))$ if and only if one of the following properties holds:*

1) $p = 1$, $s \not\leq 0$, $v_i > 0$ *for some* $i \in \mathcal{I}^{+1}(s)$, $\pi_1 = 0$, *and* $d(\pi_1) = m - 1$;

2) $1 < p < t+1$, $\pi_{p-1} = k$, $\pi_p = 0$, $k \in \mathcal{T}$, $s \leq 0$ *or* $v_i = 0$ *for all* $i \in \mathcal{I}^{+1}(s)$, *and* $d(\pi_{p-1}) = d(\pi_p) - m$;

3) $1 < p < t+1$, $\pi_{p-1} = 0$, *and* $d(\pi_{p-1}) = d(\pi_p)$;

4) $p = t+1$, *and* $d(\pi_t) = 0$;

5) $p = t+1$, σ *lies in* $A^0(s, \mathcal{T})$, $\pi_t = 0$, *and* $d(\pi_t) = m$.

When τ lies in the boundary of $A(s, \mathcal{T})$ $(A^0(s, \mathcal{T}))$ then either τ lies in $\mathbf{R}^n(s)$ or is a facet of exactly one other simplex in $A(s)$ $(A^0(s))$ or τ is itself a $(t-1)$-simplex in some $A(\overline{s})$ or $A^0(\overline{s})$ with $\overline{s}_j \neq 0$ for some $j \in \mathcal{I}^0(s)$ and $\overline{s}_h = s_h$ for $h \neq j$, depending on which case in Lemma 7.3.1 is valid.

Case 1: τ lies in $\mathbf{R}^n(s)$. If $s \not\geq 0$ and σ lies in $A(s, \mathcal{T})$ then σ shares τ with a unique t-simplex $\overline{\sigma}(\overline{y}^1, \overline{\pi})$ in $A^0(s, \mathcal{T})$ where $\overline{y}^1 = y^1 + m^{-1}q(\pi_1)$ and $\overline{\pi} = (\pi_2, \ldots, \pi_t, \pi_1)$. A

replacement step is made resulting in a t-simplex in an adjoining set $A^0(s,\mathcal{T})$. Notice that $q(0)$ is replaced by $q(0) - \sum_{j\in\mathcal{I}^{+1}(s)} v_j e(j)$.

Case 2: τ is a $(t-1)$-simplex $\overline{\sigma}(y^1,\overline{\pi})$ in $A^0(\overline{s},\overline{\mathcal{T}})$ where $\overline{s} = s + e(k)$, $\overline{\mathcal{T}} = \mathcal{T}\backslash\{k\}$, and $\overline{\pi} = (\pi_1,\ldots,\pi_{p-2},\pi_p,\ldots,\pi_t)$. Note that $q(k)$ disappears and $q(0)$ becomes $q(0) + q(k) + \sum_{j\in\mathcal{I}^{+1}(s)} v_j e(j)$ when $\sigma \in A(s,\mathcal{T})$ ($q(0) - q(k)$ when $\sigma \in A^0(s,\mathcal{T})$).

Case 3: If $\pi_p = k$ and $v_k > 0$, then σ shares τ with a t-simplex $\overline{\sigma}(y^1,\pi)$ in $A(s,\overline{\mathcal{T}})$ when $\sigma \in A(s,\mathcal{T})$ ($A^0(s,\overline{\mathcal{T}})$ when $\sigma \in A^0(s,\mathcal{T})$) where $\overline{\mathcal{T}} = \mathcal{T} \setminus \{h\} \cup \{-h\}$, $h = -k$ if $-k \in \mathcal{T}$, $h = k$ if $k \in \mathcal{T}$.

If $\pi_p = k$ and $v_k = 0$ with $-k \in \mathcal{T}$, then τ is a $(t-1)$-simplex $\overline{\sigma}(y^1,\overline{\pi})$ in $A(\overline{s},\overline{\mathcal{T}})$ when $\sigma \in A(s,\mathcal{T})$ ($A^0(\overline{s},\overline{\mathcal{T}})$ when $\sigma \in A^0(s,\mathcal{T})$) where $\overline{s} = s + e(k)$, $\overline{\mathcal{T}} = \mathcal{T}\backslash\{-k\}$, and $\overline{\pi} = (\pi_1,\ldots,\pi_{p-1},\pi_{p+1},\ldots,\pi_t)$. Here, $q(k)$ disappears and $q(0)$ becomes $q(0) + q(k)$.

Case 4: If $\pi_t = k$ and $-k \in \mathcal{T}$, then τ is a $(t-1)$-simplex $\overline{\sigma}(y^1,\overline{\pi})$ in $A(\overline{s},\overline{\mathcal{T}})$ when $\sigma \in A(s,\mathcal{T})$ ($A^0(\overline{s},\overline{\mathcal{T}})$ when $\sigma \in A^0(s,\mathcal{T})$) where $\overline{s} = s - e(k)$, $\overline{\mathcal{T}} = \mathcal{T} \setminus \{-k\}$, and $\overline{\pi} = (\pi_1,\ldots,\pi_{t-1})$.

If $\pi_t = k$ and $k \in \mathcal{T}$, then τ is a $(t-1)$-simplex $\overline{\sigma}(y^1,\overline{\pi})$ in $A(\overline{s},\overline{\mathcal{T}})$ where $\overline{s} = s + e(k)$, $\overline{\mathcal{T}} = \mathcal{T} \setminus \{k\}$, and $\overline{\pi} = (\pi_1,\ldots,\pi_{t-1})$. If $\pi_t = 0$ then $t = 1$ and so $\tau = \{v\}$.

Case 5: τ lies in $\mathbf{R}^n(s)$, and σ shares τ with a unique t-simplex $\overline{\sigma}(\overline{y}^1,\overline{\pi})$ in $A(s,\mathcal{T})$ where $\overline{y}^1 = y^1 - m^{-1}q(\pi_t)$ and $\overline{\pi} = (\pi_t,\pi_1,\ldots,\pi_{t-1})$. So, a replacement step is made resulting in a t-simplex in an adjoining set $A(s,\mathcal{T})$. Notice that $q(0)$ is replaced by $q(0) + \sum_{j\in\mathcal{I}^{+1}(s)} v_j e(j)$.

Finally, concerning Case 2 of the algorithm a t-simplex $\sigma(y^1,\pi)$ in $A(s,\mathcal{T})$ is a facet of exactly one $(t+1)$-simplex $\overline{\sigma}$ in a nonempty set $A(\overline{s})$ where $\overline{s}$ is such that $\overline{s}_k = 0$ for some $k \notin \mathcal{I}^0(s)$ while $\overline{s}_i = s_i$ for all other $i \in \mathcal{I}_n$. More precisely, $\overline{\sigma}$ is equal to $\overline{\sigma}(y^1,\overline{\pi})$ and is a $(t+1)$-simplex in $A(\overline{s},\overline{\mathcal{T}})$. If $s_k = +1$ and $v_k \neq 0$ then $\overline{\mathcal{T}} = \mathcal{T} \cup \{k\}$ and $\overline{\pi} = (\pi_1,\ldots,\pi_t,k)$. If $s_k = +1$ and $v_k = 0$ then $\overline{\mathcal{T}} = \mathcal{T} \cup \{-k\}$, $d(k) = d(0)$, and $\overline{\pi} = (\pi_1,\ldots,\pi_{p-1},k,\pi_p,\ldots,\pi_t)$ where p is such that $\pi_{p-1} = 0$. If $s_k = -1$ then $\overline{\mathcal{T}} = \mathcal{T} \cup \{-k\}$ and $\overline{\pi} = (\pi_1,\ldots,\pi_t,k)$.

A t-simplex $\sigma(y^1,\pi)$ in $A^0(s,\mathcal{T})$ is a facet of exactly one $(t+1)$-simplex $\overline{\sigma}(y^1,\overline{\pi})$ either in a nonempty set $A(\overline{s})$ or in a nonempty set $A^0(\overline{s})$. More precisely, if $s_k = +1$,

$v_k > 0$, and $v_i > 0$ for some $i \in \mathcal{I}^{+1}(\bar{s})$ then $\overline{\sigma}(y^1, \overline{\pi})$ lies in $A^0(\bar{s}, \overline{\mathcal{T}})$ where $d(k) = d(0) - m$, $\overline{\mathcal{T}} = \mathcal{T} \cup \{k\}$, and $\overline{\pi} = (\pi_1, \ldots, \pi_{p-1}, k, \pi_p, \ldots, \pi_t)$ with $\pi_p = 0$.

If $s_k = +1$, $v_k > 0$, and $v_i = 0$ for all $i \in \mathcal{I}^{+1}(\bar{s})$ then $\overline{\sigma}(y^1, \overline{\pi})$ lies in $A(\bar{s}, \overline{\mathcal{T}})$ where $d(k) = d(0) - m$, $\overline{\mathcal{T}} = \mathcal{T} \cup \{k\}$, and $\overline{\pi} = (\pi_1, \ldots, \pi_{p-1}, k, \pi_p, \ldots, \pi_t)$ with $\pi_p = 0$.

If $s_k = +1$, and $v_k = 0$ then $\overline{\sigma}(y^1, \overline{\pi})$ lies in $A^0(\bar{s}, \overline{\mathcal{T}})$ where $d(k) = d(0)$, $\overline{\mathcal{T}} = \mathcal{T} \cup \{-k\}$, and $\pi = (\pi_1, \ldots, \pi_p, k, \pi_{p+1}, \ldots, \pi_t)$ where $\pi_p = 0$.

If $s_k = -1$ then $\overline{\sigma}(y^1, \overline{\pi})$ lies in $A^0(\bar{s}, \overline{\mathcal{T}})$ where $\overline{\mathcal{T}} = \mathcal{T} \cup \{k\}$, $\overline{\pi} = (\pi_1, \ldots, \pi_t, k)$.

This concludes the description of how to follow a sequence of adjacent simplices in the simplicial subdivision $\mathcal{G}_m$ of $\mathbf{R}^n_+$.

Part C

An economy with linear production technologies

Chapter 8

SLCP-algorithms to compute an economic equilibrium

Given some economic model one is often interested in computing a situation where the economy under consideration is said to be in equilibrium. An equilibrium in an economy is a state in which no individual taking part in this economy is eager to change behaviour. Often such an equilibrium takes the form of one of the complementarity problems introduced in Part B of this monograph.

The most simple economic model is a pure exchange economy. Such an economy consists of agents who are endowed with some initial bundle of commodities and have a preference relation over these commodities represented by a utility function. Each commodity can be exchanged with another commodity at a certain rate, the price rate of these commodities. Given the prices of the commodities every agent in the economy expresses a demand or supply of each commodity maximizing utility. Finding an equilibrium in this economy is then equivalent to computing prices of commodities such that the total demand of the agents does not exceed the amount of commodities supplied by the agents in this economy. In Doup (1988) several algorithms were introduced which are capable of computing an approximating equilibrium in a pure exchange economy for any finite accuracy. These algorithms are path following algorithms where the path can be interpreted as a tâtonnement process in which the prices are simultaneously adjusted (compare Walras (1874)).

The next step is to introduce the possibility of production in this exchange econ-

omy. One of the main contributions in this field is the computation of an equilibrium in an economy with so-called linear production technologies by the algorithms introduced in Mathiesen (1985b) and Eaves (1987). They proposed to solve the equilibrium problem in an economy with linear production technologies by approximating the equilibrium conditions by a sequence of linear complementarity problems and solving each linear complementarity problem by applying the Lemke complementarity pivoting algorithm. In this way a sequence of prices may be generated possibly converging to an equilibrium price vector.

The drawback of Mathiesen's algorithm is the possible breakdown in an iteration of the sequence of linear complementarity problems when the Lemke complementary pivoting algorithm fails to compute a solution to the linear complementarity problem. Eaves on the other hand manages to generate a sequence of prices although at the cost of solving a high dimensional linear complementarity problem in each iteration of the sequence. In Chapter 9 of this monograph we introduce an algorithm which efficiently escapes these problems. As an alternative to the algorithm introduced in Chapter 9 we introduce a simplicial algorithm to solve the equilibrium problem in an economy with linear production technologies in Chapter 10. Contrary to the algorithms introduced in Mathiesen (1985b), Eaves (1987), and Chapter 9, the path followed by the simplicial algorithm has an interpretation of a tâtonnement process. This tâtonnement process has been described in van den Elzen, van der Laan, and Talman (1990). The present chapter describes an economy with linear production technologies in Section 1. Given the equilibrium conditions resulting from the description in Section 1 we review the algorithm introduced in Mathiesen (1985b) in Section 2 and the algorithm introduced in Eaves (1987) in Section 3.

8.1 The economy with linear production technologies

Consider an economy with a finite number of consumers, commodities, and production activities. Each consumer in the economy is assumed to have an initial endowment in each of the, say $n+1$, commodities. Let ω denote the $(n+1)$-vector of total initial endowments where ω_j is the total endowment in commodity j, $j \in \mathcal{I}_{n+1}$. Given

a price vector $p \in \mathbf{R}_+^{n+1} \setminus \{0\}$ with p_j denoting the price of commodity j, let $d(p)$ denote the total demand of the consumers where $d_j(p)$ is the demand for commodity $j \in \mathcal{I}_{n+1}$.

Assumption 8.1.1 *The demand function d has the following properties:*

i) d is continuous in $p \in \mathbf{R}_+^{n+1} \setminus \{0\}$;

ii) d is homogeneous of degree zero, i.e., $d(\lambda p) = d(p)$ for all $\lambda > 0$ and $p \in \mathbf{R}_+^{n+1} \setminus \{0\}$;

iii) $p^\top d(p) = p^\top \omega$ for every $p \in \mathbf{R}_+^{n+1} \setminus \{0\}$.

At price vector $p \in \mathbf{R}_+^{n+1} \setminus \{0\}$ the excess demand is denoted by $z(p)$, i.e. $z_j(p)$ equals $d_j(p) - \omega_j$. The properties of the excess demand function follow immediately from Assumption 8.1.1. For an economic interpretation of these properties we refer to Debreu (1959).

Property 8.1.1 *The excess demand function z has the following properties:*

i) z is continuous in $p \in \mathbf{R}_+^{n+1} \setminus \{0\}$;

ii) z is homogeneous of degree zero;

iii) (Walras' law) $p^\top z(p) = 0$ for every $p \in \mathbf{R}_+^{n+1} \setminus \{0\}$.

Commodities in the economy can be produced by a finite number of, say l, activities. Activity j, $j \in \mathcal{I}_l$, is represented by a vector $a^j \in \mathbf{R}^{n+1}$. The column vector a^j has components a_s^j such that $a_s^j \geq 0$ implies that activity j has an output of a_s^j units of commodity s while $a_s^j \leq 0$ implies that activity j uses $-a_s^j$ units of commodity s as an input, when the activity level equals one. The level of activity j is denoted by the nonnegative number y_j. If one puts all activity levels y_j into an l-vector y and all columns a^j into an $(n+1) \times l$-matrix A then the vector Ay denotes the aggregate net input-output for all commodities in the economy given the activity level vector $y \in \mathbf{R}_+^l$. So, given $y \geq 0$, $(Ay)_s > 0$ implies that commodity s serves as a net output to the economy as a whole and $(Ay)_s < 0$ implies that commodity s serves as a net input to the economy as a whole. With respect to the activity matrix A we assume that in case of production there exists at least one commodity serving as an input to the economy. Formally, we have the following assumption.

Assumption 8.1.2 *If $Ay \geq 0$ and $y \geq 0$ then $y = 0$.*

We assume that the producers running the activities in the economy choose the activity levels in such a way that the profits made by each activity are maximized. If the profit $p^T a^j$ for some activity j is negative then the producer running activity j sets the activity level y_j equal to zero. Raising y_j from zero would bring losses to the producer. If $p^T a^j > 0$ then y_j is raised towards infinity by the producer running activity j while if $p^T a^j = 0$ then any activity level y_j can be chosen such that $y_j \geq 0$.

The economy described above is said to be in equilibrium if the prices of the commodities in the economy and the activity levels are such that for every commodity total demand of the consumers is met by the initial endowment of the economy in this commodity plus its total net input-output of the activities. This condition implies that in equilibrium no activity makes positive profit. Therefore the following definition constitutes an equilibrium in an economy with linear production technologies.

Definition 8.1.1 *An equilibrium in an economy with linear production technologies is a price vector $p^* \in \mathbf{R}^{n+1}_+ \setminus \{0\}$ and a vector with activity levels $y^* \geq 0$ such that*

i) $z(p^*) - Ay^* \leq 0$: market-clearance;

ii) $p^{*T} A \leq 0$: no profits.

The properties of the market excess demand function together with the equilibrium conditions on p and y imply the following relations between the equilibrium prices and activity levels in the economy.

Property 8.1.2 *If a price vector p^* and an activity level vector y^* constitute an equilibrium in an economy with linear production technologies then*

$$p^{*T} A y^* = 0.$$

Proof: From the market-clearance condition in Definition 8.1.1 it follows that $p^{*T}(z(p^*) - Ay^*) = p^{*T} z(p^*) - p^{*T} Ay^* = -p^{*T} Ay^* \leq 0$, applying Walras' law. Furthermore, multiplying $p^{*T} A \leq 0$ by $y^* \geq 0$ gives $p^{*T} Ay^* \leq 0$. Together with $-p^{*T} Ay^* \leq 0$ this implies that $p^{*T} Ay^* = 0$. □

Property 8.1.2 says that in equilibrium an activity showing a deficit ($p^{*\mathsf{T}}a^j < 0$) is not producing ($y_j^* = 0$) while an activity in operation ($y_j^* > 0$) runs at balance ($p^{*\mathsf{T}}a^j = 0$). This also follows from the presumed profit maximizing behaviour of the producers running the activities.

Property 8.1.3 *If a price vector p^* and a vector of activity levels y^* constitute an equilibrium in an economy with linear production technologies then*

$$p^{*\mathsf{T}}(z(p^*) - Ay^*) = 0.$$

Proof: $p^{*\mathsf{T}}(z(p^*) - Ay^*) = p^{*\mathsf{T}}z(p^*) - p^{*\mathsf{T}}Ay^* = 0$, by applying Walras' law and Property 8.1.2. □

Property 8.1.3 says that in equilibrium a commodity in excess supply ($z_i(p^*) - (Ay^*)_i < 0$) has a price equal to zero ($p_i^* = 0$) while a commodity with positive price ($p_i^* > 0$) implies market clearance ($z_i(p^*) - (Ay^*)_i = 0$).

8.2 Mathiesen's *SLCP*-algorithm

In Mathiesen (1985b) the following definition of an equilibrium in an economy with linear production technologies is used.

> A price vector p^* and a vector of activity levels y^* constitute an equilibrium in an economy with linear production technologies if
>
> $$\begin{aligned} &z(p^*) - Ay^* \leq 0 \\ &p^{*\mathsf{T}}A \leq 0 \\ &p^{*\mathsf{T}}Ay^* = 0 \\ &p^{*\mathsf{T}}(z(p^*) - Ay^*) = 0 \\ &p^* \geq 0,\ y^* \geq 0. \end{aligned} \tag{8.1}$$

The reason for defining an equilibrium in this way and not in the way of Definition 8.1.1 is that this form directly fits into what Mathiesen calls the "complementarity format". If one defines the mapping f from $\mathbf{R}_+^{n+1+l}$ into $\mathbf{R}^{n+1+l}$ as

$$f\begin{pmatrix} p \\ y \end{pmatrix} = \begin{pmatrix} -z(p) + Ay \\ -A^{\mathsf{T}}p \end{pmatrix}$$

and $x \in \mathbf{R}_+^{n+1+l}$ as $x^\top = (p^\top, y^\top)$ then an equilibrium in an economy with linear production technologies corresponds to an x^* such that

$$f(x^*) \geq 0,\ x^* \geq 0,\ x^{*\top} f(x^*) = 0,$$

which is exactly the definition of a nonlinear complementarity problem as defined in Chapter 7 of this monograph.

An equilibrium in an economy with linear production technologies as defined in (8.1) cannot be computed directly. We have to take notice of the fact that because of homogeneity of degree zero of the excess demand function in the prices (Property 8.1.1ii) the existence of an equilibrium price vector p^* implies the existence of a ray of equilibrium price vectors $\{\lambda p^* \mid \lambda > 0\}$. To overcome this problem the prices should be normalized.

To compute an equilibrium in an economy with linear production technologies Mathiesen approximates the nonlinear complementarity problem by a sequence of linear complementarity problems (*SLCP*). This sequence of linear complementarity problems consists of a sequence of iterates where in each iterate the nonlinear complementarity problem is linearized in a price vector either obtained as a solution to the previous iterate in the sequence or, if no previous iterate exists, chosen arbitrarily.

Consider iteration k of the *SLCP*. In this iteration the algorithm linearizes the nonlinear complementarity problem by replacing the nonlinear excess demand function z with its first-order Taylor expansion in the price vector obtained as a solution to the previous iterate in case $k \geq 1$ and chosen arbitrarily from S^n in case $k = 0$. Denote this price vector with p^k. Without loss of generality, p^k lies in the n-dimensional unit-simplex S^n, so $e^\top p^k = 1$. We assume that the excess demand function is differentiable in each point $p \in \mathbf{R}_+^{n+1} \setminus \{0\}$. Then the first-order Taylor expansion of z in p^k, denoted z^k, exists and is equal to

$$z^k(p) = z(p^k) + \nabla z(p^k)p$$

for all $p \in \mathbf{R}_+^{n+1} \setminus \{0\}$. Notice that due to the homogeneity of z we have that $\nabla z(p^k)p^k = 0$.

Replacing z by its first-order Taylor expansion z^k in p^k in iteration k the problem would become to find a vector $p \in \mathbf{R}^{n+1}$ and a vector $y \in \mathbf{R}^l$ such that

$$\begin{aligned} & z(p^k) + \nabla z(p^k)p - Ay \leq 0 \\ & p^\top A \leq 0 \\ & p^\top Ay = 0 \\ & p^\top (z(p^k) + \nabla z(p^k)p - Ay) = 0 \\ & p \geq 0,\ y \geq 0. \end{aligned} \tag{8.2}$$

The property of the Jacobian matrix $\nabla z(p)$ that $\nabla z(p)p = 0$ for all $p \in \mathbf{R}^{n+1}_+ \setminus \{0\}$ causes the Jacobian matrix of first-order derivatives to be singular. Moreover, system (8.2) might not have a solution. Hence system (8.2) is inappropriate. This impasse is traditionally escaped either by adding a normalization constraint such as $e^\top p = 1$ to (8.2) or by stipulating a numeraire commodity having fixed price. In Mathiesen (1985b) the latter approach is chosen.

Let commodity m be chosen as a numeraire. Then we can stipulate that its price is fixed at $p_m = p^k_m$. According to Mathiesen (1985b) this gives the following linear complementarity problem to solve in iteration k

$$\begin{aligned} & \hat{z}(p^k) + \nabla \hat{z}(p^k)\hat{p} - \hat{A}y \leq 0 \\ & \hat{p}^\top \hat{A} \leq 0 \\ & \hat{p}^\top \hat{A}y = 0 \\ & \hat{p}^\top (\hat{z}(p^k) + \nabla \hat{z}(p^k)\hat{p} - \hat{A}y) = 0 \\ & \hat{p} \geq 0,\ y \geq 0, \end{aligned} \tag{8.3}$$

where $\hat{p}$ denotes the price vector p without the numeraire price p_m in p^k, the n-vector $\hat{z}(p^k)$ denotes the vector with the excess demand in all commodities except the m-th commodity, the matrix $\nabla \hat{z}(p^k)$ denotes the matrix with first-order derivatives of z in p^k, $\nabla z(p^k)$, with the m-th row and the m-th column deleted, and the matrix $\hat{A}$ is the matrix A with the m-th row deleted. Hence, (8.3) is obtained by fixing p_m and dropping the m-th commodity's linearized balance from (8.2). Notice that Walras' law does not hold for z^k except in the starting point of the iteration. If one succeeds in computing a solution to the linear complementarity problem in (8.3) then one ends up with a solution where the m-th commodity linearized balance as well as the per unit profit made by the m-th commodity is possibly out of equilibrium. Furthermore

there are $n+1$ commodities which can be chosen to be a numeraire. This gives the possibility to choose from $n+1$ possible linear complementarity problems in each iteration of the *SLCP*.

Mathiesen (1985b) attempts to solve the obtained linear complementarity problem in each iteration by applying the Lemke complementary pivoting algorithm. In Chapter 4 we saw that the algorithm either solves the linear complementarity problem or ends up in a ray indicating that no solution exists to the linear complementarity problem. If the Lemke complementary pivoting algorithm computes a solution to the linear complementarity problem in (8.3) this solution is taken as a starting point to the next iteration. Notice however that the computed solution may not lie in S^n. Then the obtained solution, say $\overline{p}$, is projected upon S^n by taking $p^{k+1} = (e^\top \overline{p})^{-1}\overline{p}$ as the starting point to the $(k+1)$-th iteration. In case the obtained solution lies in the boundary of S^n, i.e. there exist prices equal to zero, the properties of the demand function might imply that the excess demand for these commodities diverge to infinity. In such a case Mathiesen proposes to modify the obtained solution $\overline{p}$ to

$$p^{k+1} = p^k + \theta(\overline{p} - p^k), \text{ for some } \theta,\ 0 < \theta < 1.$$

In many examples given in Mathiesen (1987) this modification of the starting point causes the algorithm to zigzag along the boundary of S^n.

If the algorithm ends up with a ray to the linear complementarity problem, Mathiesen proposes to choose another commodity as a numeraire and to repeat the same procedure for the obtained linear complementarity problem. If in this way the algorithm is able to compute a solution in each iteration of the *SLCP* a sequence of prices $\{p^k\}_{k=0}^{\infty}$ is generated possibly converging towards an equilibrium price vector. Mathiesen is not able to show global convergence of his *SLCP*-algorithm although he claims that his experiments support global convergence. Notice that Mathiesen also gives a counter example to possible global convergence of his *SLCP*-algorithm, namely Scarf's unstable exchange equilibrium models (see Scarf (1960)).

In case the *SLCP* contains an iteration in which there exists no linear complementarity problem for which a solution can be calculated by the Lemke complementary pivoting algorithm, the *SLCP* will break down and no sequence of prices can be generated. Mathiesen is not able to show local convergence in each iterate of the sequence of linear complementarity problems either. He only claims local convergence on the

basis of some experiments in Mathiesen (1985a), Mathiesen (1985b), and Mathiesen and Rutherford (1983).

It may be clear that the *SLCP* as proposed in Mathiesen (1985b) contains several drawbacks. Neither global nor local convergence can be proved. Furthermore in each iteration one has to choose among $n+1$ possible linear complementarity problems to solve. If the Lemke complementary pivoting algorithm is not able to compute a solution to one of these linear complementarity problems Mathiesen just chooses to solve one of the other possible linear complementarity problems. This makes his *SLCP*-algorithm very close to a "trial-and-error"-method. Apart from these drawbacks the algorithm as proposed by Mathiesen (1985b) does not process the information obtained from the previous iteration in a very efficient way. This is caused by the fact that the Lemke complementary pivoting algorithm is stuck to a fixed starting point. To remove this drawback one can apply in each iterate the algorithm introduced in Chapter 5 to solve the linear complementarity problem. Finally, the algorithm does not take into account any information about the numeraire commodity so that at the solution $\overline{p}$ the linearized balance as well as the per unit profit for this commodity might be far out of equilibrium.

8.3 Eaves' *SLCP*-algorithm

The impossibility to prove local as well as global convergence of the sequence of linear complementarity problems as introduced in Mathiesen (1985b) motivated Eaves to introduce an alternative *SLCP* which should lend itself more easily to theoretical understanding. Furthermore Eaves' *SLCP* converges in each iteration of the sequence thereby assuring the generation of a sequence of prices in S^n, possibly converging to an equilibrium price vector.

Eaves proves that the problem to compute an equilibrium in an economy with linear production technologies is equivalent to the stationary point problem of the excess demand function z on a set S_A^n defined as $S_A^n = \{p \in S^n \mid p^\top A \leq 0\}$.

Theorem 8.3.1 *A price vector $p^* \in S_A^n$ is an equilibrium price vector in an economy with linear production technologies if and only if p^* is a stationary point of the excess demand function z on S_A^n.*

Proof: Let $p^* \in S_A^n$ and $y^* \geq 0$ constitute an equilibrium in an economy with linear production technologies. Then

$$p^\top z(p^*) \leq p^\top z(p^*) - p^\top A y^* = p^\top (z(p^*) - A y^*) \leq 0 = p^{*\top} z(p^*)$$

for all $p \in S_A^n$. Therefore $p^* \in S_A^n$ is a stationary point of z on S_A^n.

Let $p^* \in S_A^n$ be a stationary point of z on S_A^n. Then for all $p \in S_A^n$ it holds that $p^\top z(p^*) \leq p^{*\top} z(p^*) = 0$. This is equivalent to p^* solving the optimization problem given by

$$\begin{array}{ll} \max & p^\top z(p^*) \\ \text{s.t.} & p^\top A \leq 0 \\ & e^\top p = 1 \\ & p \geq 0. \end{array} \tag{8.4}$$

Take the vector y as the vector with the dual variables to the constraints in $p^\top A \leq 0$, β as the dual variable to the constraint $e^\top p = 1$, and $\mu \geq 0$ as the vector with dual variables to the constraints in $p \geq 0$. Then p^* being a solution to the optimization problem above implies that p^* fulfils the following first-order conditions:

$$\begin{array}{rcl} z(p^*) & = & Ay + \beta e - \mu \\ A^\top p^* & \leq & 0 \\ e^\top p^* & = & 1 \\ \mu^\top p^* & = & 0,\ p^{*\top} A y = 0 \\ \mu & \geq & 0,\ y \geq 0,\ \beta \in \mathbf{R}. \end{array} \tag{8.5}$$

Premultiplying $z(p^*) = Ay + \beta e - \mu$ by p^* results in $\beta = 0$. Thus, with $\mu \geq 0$ it follows that $z(p^*) - Ay = -\mu \leq 0$. Therefore p^* is an equilibrium price vector in an economy with linear production technologies. □

In order to solve this stationary point problem Eaves suggests to approximate this nonlinear problem by a sequence of linear stationary point problems. Each solution to the linear stationary point problem in this sequence is a solution to the linear stationary point problem which arises when replacing the excess demand function z in the economy by its first-order Taylor expansion z^k in the price vector obtained as a solution to the previous iteration or, if no previous iteration exists, chosen arbitrarily from S_A^n.

Consider iteration k of the *SLCP*. In iteration k of the *SLCP* the algorithm solves the stationary point problem of z^k on S_A^n by generating a piecewise linear path of points in S_A^n starting in p^k and ending up with a solution p^{k+1} which will be used as the starting point to the next iteration. Each point on the path can be seen as a stationary point of z^k on the set $S_A^n \cap S^n(\theta)$ for some θ, $0 \leq \theta \leq 1$. The set $S^n(\theta)$, $0 \leq \theta \leq 1$, is defined as

$$S^n(\theta) = (1-\theta)p^k + \theta S^n.$$

The stationary point problem on the parametrized set $S_A^n \cap S^n(\theta)$ can be interpreted as a homotopy problem with θ as the homotopy parameter, θ running from zero to one. If $\theta = 0$ then $S_A^n \cap S^n(0) = S^n(0) = \{p^k\}$ and the starting point p^k is a trivial solution to the stationary point problem of z^k on $S_A^n \cap S^n(0)$. If θ becomes equal to one in a point $\overline{p}$ on the path generated by the algorithm then $\overline{p}$ is a stationary point of z^k on $S_A^n \cap S^n(1) = S_A^n$ and hence a solution to the problem in iteration k. Then $p^{k+1} = \overline{p}$ is taken as the starting point to the next iteration. In Figure 8.3.1 the set $S_A^n \cap S^n(\theta)$ is illustrated for three values of θ for $n = 2$ and $l = 2$.

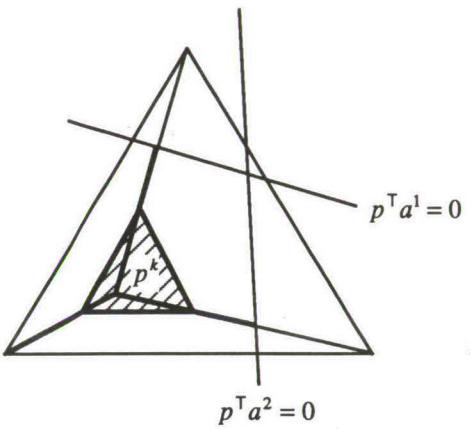
A point $\overline{p}$ is a stationary point of z^k on $S_A^n \cap S^n(\overline{\theta})$ for some $\overline{\theta} \geq 0$ if $\overline{p}$ maximizes $p^\top z^k(\overline{p})$ over $p \in S_A^n \cap S^n(\overline{\theta})$. Then by definition of $S_A^n \cap S^n(\overline{\theta})$ there exists an $(n+1)$-vector $\overline{\lambda}$ such that $\overline{p} = (1-\overline{\theta})p^k + \overline{\lambda}$ and $\overline{\lambda}$ solves the maximization problem, denoted as the primal, given by

$$\begin{aligned} \max\ & \lambda^\top z^k(\overline{p}) \\ \text{s.t.}\ & \lambda \geq 0 \\ & e^\top \lambda = \overline{\theta} \\ & A^\top \lambda \leq -(1-\overline{\theta})A^\top p^k. \end{aligned}$$

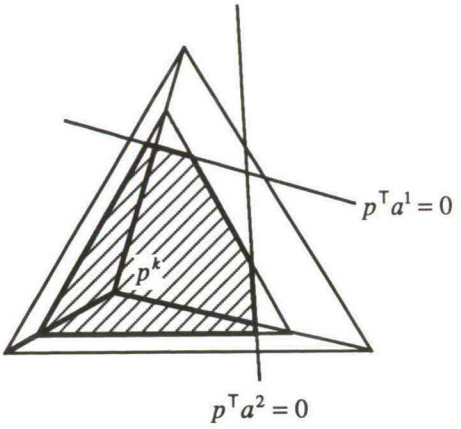
This is a linear programming problem. According to the Duality Theorem of Linear Programming, solving this maximization problem is equivalent to solving its dual

$$\begin{aligned} \min\ & \beta\overline{\theta} - (1-\overline{\theta})p^{k\top} Ay \\ \text{s.t.}\ & Ay - \mu + \beta e = z^k(\overline{p}) \\ & y \geq 0,\ \mu \geq 0,\ \beta \in \mathbf{R}, \end{aligned}$$

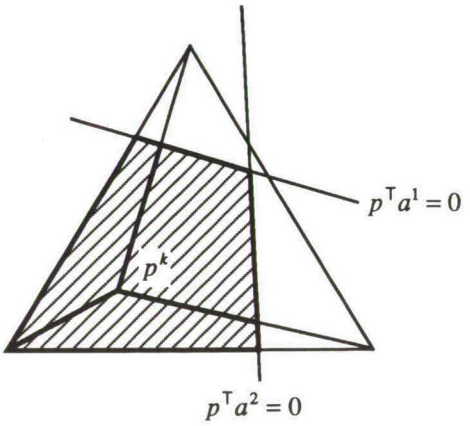
where y denotes the l-vector with dual variables to the constraints in $A^\top \lambda \leq -(1-\overline{\theta})A^\top p^k$, as its components, μ denotes the $(n+1)$-vector with the dual variables to the constraints in $\lambda \geq 0$ as its components, and β denotes the dual variable to $e^\top \lambda = \overline{\theta}$.



(a) $S_A^2 \cap S^2(\frac{1}{3}) = S^2(\frac{1}{3})$



(b) $S_A^2 \cap S^2(\frac{2}{3})$



(c) $S_A^2 \cap S^2(1) = S_A^2$

FIGURE 8.3.1: The subset $S_A^2 \cap S^2(\theta)$ for $\theta = \frac{1}{3}, \frac{2}{3}, 1$ with $p^k = (\frac{3}{5}, \frac{1}{5}, \frac{1}{5})^T$ in an economy with two linear production technologies and three commodities.

If $\overline{\lambda} = \overline{p} - (1 - \overline{\theta})p^k$ solves the primal then there exist a unique $\overline{\mu} \geq 0$, $\overline{y} \geq 0$, and $\overline{\beta} \in \mathbf{R}$ solving the dual. Let $\overline{v} = A^T\overline{p}$. If we let v denote the l-vector with slack variables to $A^T\lambda \leq -(1-\theta)A^Tp^k$ then $\overline{\lambda}$, $\overline{\mu}$, $\overline{y}$, $\overline{\theta}$, $\overline{\beta}$, and $\overline{v}$ solve

$$\begin{array}{rcccccccc} \mu & + & \nabla z(p^k)\lambda & - & Ay & - & \beta e & = & -z(p^k) \\ v & + & A^T\lambda & & & - & A^Tp^k\theta & = & -A^Tp^k \\ & & e^T\lambda & & & - & \theta & = & 0 \end{array} \tag{8.6}$$

$$\mu^T\lambda = 0,\ v^Ty = 0$$

$$\mu \geq 0,\ \lambda \geq 0,\ v \geq 0,\ y \geq 0,\ 0 \leq \theta \leq 1.$$

On the other hand, if a solution $\overline{\lambda}$, $\overline{\mu}$, $\overline{y}$, $\overline{\beta}$ to (8.6) for some $\overline{\theta}$, $0 \leq \overline{\theta} \leq 1$, is found then $\overline{\lambda}$ solves the primal and the point $\overline{p} = (1 - \overline{\theta})p^k + \overline{\lambda}$ is a stationary point of z^k on $S_A^n \cap S^n(\overline{\theta})$. Hence $\overline{p}$ is a point on the path generated by the algorithm. If $\overline{\theta} = 1$ then a solution to the stationary point problem in iteration k is found.

The algorithm generates a path of points starting in p^k towards a solution of the stationary point problem of z^k on S_A^n by generating solutions λ, μ, y, θ, β to (8.6) for θ, $0 \leq \theta \leq 1$. This means that, in iteration k, the algorithm attempts to solve the linear stationary point problem $LSPP_k$ given by

$$\begin{array}{rcccccccc} \mu & + & \nabla z(p^k)\lambda & - & Ay & - & \beta e & = & -z(p^k) \\ v & + & A^T\lambda & & & & & = & 0 \\ & & e^T\lambda & & & & & = & 1 \end{array} \tag{8.7}$$

$$\mu^T\lambda = 0,\ v^Ty = 0$$

$$\mu \geq 0,\ \lambda \geq 0,\ v \geq 0,\ y \geq 0,\ \beta \in \mathbf{R}.$$

Eaves suggests to compute a solution to $LSPP_k$ by applying the Lemke complementary pivoting algorithm in a slightly different way as introduced in Chapter 4 of this monograph. The difference is however only a matter of redefinition of the variables.

The algorithm of Eaves starts with p^k as an end point of a line segment of solutions $\lambda_r \geq 0$, $v_j \geq 0$ $(j \in \mathcal{I}_l)$, $\mu_j \geq 0$ $(j \in \mathcal{I}_{n+1} \setminus \{r\})$, $0 \leq \theta \leq 1$ to the system of equations given by

$$\lambda_r \begin{pmatrix} \nabla z_r(p^k) \\ (A^T)_{.r} \\ 1 \end{pmatrix} + \sum_{j=1}^{l} v_j \begin{pmatrix} 0 \\ e(j) \\ 0 \end{pmatrix} + \sum_{j=1,\ j\neq r}^{n+1} \mu_j \begin{pmatrix} e(j) \\ 0 \\ 0 \end{pmatrix} - \beta \begin{pmatrix} e \\ 0 \\ 0 \end{pmatrix} - \theta \begin{pmatrix} 0 \\ A^Tp^k \\ 1 \end{pmatrix} = \begin{pmatrix} -z(p^k) \\ -A^Tp^k \\ 0 \end{pmatrix}$$

for some $r \in \mathcal{I}_{n+1}$, assuming nondegeneracy. Notice that the assumption of non-degeneracy on p^k implies that $p^{k\top}A < 0$. If there exists some $t \in \mathcal{I}_l$ such that $p^{k\top}a^t = 0$ then it can easily be seen that p^k is also a line segment of solutions $y_t \geq 0$, $v_j \geq 0\ (j \in \mathcal{I}_l \backslash \{t\})$, $\mu_j \geq 0\ (j \in \mathcal{I}_{n+1})$, $0 \leq \theta \leq 1$ to the system of equations given by

$$-y_t \begin{pmatrix} a^t \\ 0 \\ 0 \end{pmatrix} + \sum_{j=1,\ j\neq t}^{l} v_j \begin{pmatrix} 0 \\ e(j) \\ 0 \end{pmatrix} + \sum_{j=1}^{n+1} \mu_j \begin{pmatrix} e(j) \\ 0 \\ 0 \end{pmatrix} - \beta \begin{pmatrix} e \\ 0 \\ 0 \end{pmatrix} - \theta \begin{pmatrix} 0 \\ A^\top p^k \\ 1 \end{pmatrix} = \begin{pmatrix} -z(p^k) \\ -A^\top p^k \\ 0 \end{pmatrix}$$

and hence p^k is a degenerate end point on the path.

In general Eaves' algorithm generates a line segment of $(\mathcal{G}^1, \mathcal{G}^2)$-complete points in subsets $A(\mathcal{G}^1) \times A(\mathcal{G}^2)$ of $\mathbf{R}^{n+1+l}$ given by the system of $n+l+1$ equations in $n+l+2$ variables

$$\begin{aligned} &\textstyle\sum_{j\in\mathcal{G}^1} \lambda_j \begin{pmatrix} \nabla z_j(p^k) \\ (A^\top)_{\cdot j} \\ 1 \end{pmatrix} + \sum_{j\in\mathcal{G}^2} v_j \begin{pmatrix} 0 \\ e(j) \\ 0 \end{pmatrix} + \sum_{j\notin\mathcal{G}^1} \mu_j \begin{pmatrix} e(j) \\ 0 \\ 0 \end{pmatrix} - \\ &\textstyle\sum_{j\notin\mathcal{G}^2} y_j \begin{pmatrix} a^j \\ 0 \\ 0 \end{pmatrix} - \beta \begin{pmatrix} e \\ 0 \\ 0 \end{pmatrix} - \theta \begin{pmatrix} 0 \\ A^\top p^k \\ 1 \end{pmatrix} = \begin{pmatrix} -z(p^k) \\ -A^\top p^k \\ 0 \end{pmatrix}, \end{aligned} \tag{8.8}$$

where $\mathcal{G}^1$ denotes the set $\{j \mid \mu_j = 0\}$ and $\mathcal{G}^2$ denotes the set $\{j \mid y_j = 0\}$, for all points on the line segment and where $A(\mathcal{G}^1)$ in $\mathbf{R}^{n+1}$ and $A(\mathcal{G}^2)$ in $\mathbf{R}^l$ are defined as in Definition 4.2.2. If in an end point of a line segment of solutions to (8.5) θ becomes equal to one then this end point contains a solution to $LSPP_k$.

Eaves claims that in each iteration of the sequence of linear complementarity problems a solution to $LCP(q, M)$ will be found, i.e. the Lemke complementary pivoting algorithm can not generate a ray of solutions to (8.5) for any sets $\mathcal{G}^1 \subset \mathcal{I}_{n+1}$ and $\mathcal{G}^2 \subseteq \mathcal{I}_l$. This is proved in the next theorem.

Theorem 8.3.2 *For all possible sets* $\mathcal{G}^1 \subset \mathcal{I}_{n+1}$ *and* $\mathcal{G}^2 \subseteq \mathcal{I}_l$ *and all* $\hat{p} \in S_A^n$ *the system of equations*

$$\sum_{j\in\mathcal{G}^1} \lambda_j \begin{pmatrix} \nabla z_j(\hat{p}) \\ (A^\top)_{\cdot j} \\ 1 \end{pmatrix} + \sum_{j\in\mathcal{G}^2} v_j \begin{pmatrix} 0 \\ e(j) \\ 0 \end{pmatrix} + \sum_{j\notin\mathcal{G}^1} \mu_j \begin{pmatrix} e(j) \\ 0 \\ 0 \end{pmatrix} -$$

$$\sum_{j\notin\mathcal{G}^2} y_j \begin{pmatrix} a^j \\ 0 \\ 0 \end{pmatrix} - \beta \begin{pmatrix} e \\ 0 \\ 0 \end{pmatrix} - \theta \begin{pmatrix} 0 \\ A^\top \hat{p} \\ 1 \end{pmatrix} = \begin{pmatrix} -z(\hat{p}) \\ -A^\top \hat{p} \\ 0 \end{pmatrix} \tag{8.9}$$

does not contain a ray of solutions satisfying $\lambda_j \geq 0$ $(j \in \mathcal{G}^1)$, $v_j \geq 0$ $(j \in \mathcal{G}^2)$, $\mu_j \geq 0$ $(j \notin \mathcal{G}^1)$, $y_j \geq 0$ $(j \notin \mathcal{G}^2)$, $\beta \in \mathbf{R}$, *and* $0 \leq \theta \leq 1$.

Proof: Suppose (8.8) contains a ray of admissable solutions $\lambda_j^0 + \alpha\lambda_j^1$ $(j \in \mathcal{G}^1)$, $v_j^0 + \alpha v_j^1$ $(j \in \mathcal{G}^2)$, $\mu_j^0 + \alpha\mu_j^1$ $(j \notin \mathcal{G}^1)$, $y_j^0 + \alpha y_j^1$ $(j \notin \mathcal{G}^2)$, $\beta^0 + \alpha\beta^1$, and $\theta^0 + \alpha\theta^1$ for $\alpha \geq 0$. Then λ_j^0 $(j \in \mathcal{G}^1)$, v_j^0 $(j \in \mathcal{G}^2)$, μ_j^0 $(j \notin \mathcal{G}^1)$, y_j^0 $(j \notin \mathcal{G}^2)$, β^0, and θ^0 is an admissable solution to (8.8) and λ_j^1 $(j \in \mathcal{G}^1)$, v_j^1 $(j \in \mathcal{G}^2)$, μ_j^1 $(j \notin \mathcal{G}^1)$, y_j^1 $(j \notin \mathcal{G}^2)$, β^1, and θ^1 is a solution to the homogeneous system of equations

$$\sum_{j\in\mathcal{G}^1} \lambda_j \begin{pmatrix} \nabla z_j(\hat{p}) \\ (A^\top)_{\cdot j} \\ 1 \end{pmatrix} + \sum_{j\in\mathcal{G}^2} v_j \begin{pmatrix} 0 \\ e(j) \\ 0 \end{pmatrix} + \sum_{j\notin\mathcal{G}^1} \mu_j \begin{pmatrix} e(j) \\ 0 \\ 0 \end{pmatrix} -$$

$$\sum_{j\notin\mathcal{G}^2} y_j \begin{pmatrix} a^j \\ 0 \\ 0 \end{pmatrix} - \beta \begin{pmatrix} e \\ 0 \\ 0 \end{pmatrix} - \theta \begin{pmatrix} 0 \\ A^\top \hat{p} \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \tag{8.10}$$

It is obvious that $\theta^1 = 0$ as $\theta^1 \neq 0$ would imply that there exists some $\alpha > 0$ such that $\theta^0 + \alpha\theta^1 > 1$ or $\theta^0 + \alpha\theta^1 < 0$. If $\lambda_h^1 < 0$ for some $h \in \mathcal{G}^1$ then there exists an $\alpha > 0$ such that $\lambda_h^0 + \alpha\lambda_h^1 < 0$ which is a contradiction. Hence $\lambda_j^1 \geq 0$ $(j \in \mathcal{G}^1)$, and similarly $v_j^1 \geq 0$ $(j \in \mathcal{G}^2)$. Then $\sum_{j\in\mathcal{G}^1} \lambda_j^1 = 0$ implies that $\lambda_j^1 = 0$ for all $j \in \mathcal{G}^1$. The second set of equations in (8.9) then reduces to $\sum_{j\in\mathcal{G}^2} v_j^1 e(j) = 0$ implying $v_j^1 = 0$ for

all $j \in \mathcal{G}^2$. If we premultiply the remaining system of equations with $p = (1-\theta^0)\hat{p}+\lambda^0$ we obtain $\beta^1 = 0$ since $\sum_{j\notin\mathcal{G}^2} p^\top a^j y_j^1 = 0$. Then it follows that

$$\sum_{j\notin\mathcal{G}^1} \mu_j^1 e(j) = \sum_{j\notin\mathcal{G}^2} y_j^1 a^j. \tag{8.11}$$

Notice that $y_j^1 \geq 0$ $(j \notin \mathcal{G}^2)$ and $\mu_j^1 \geq 0$ $(j \notin \mathcal{G}^1)$. This follows from $y_j^0 + \alpha y_j^1$ $(j \notin \mathcal{G}^2)$ and $\mu_j^0 + \alpha\mu_j^1$ $(j \notin \mathcal{G}^1)$ being a solution to (8.8) for all $\alpha \geq 0$. Indeed, $y_h^1 < 0$ for some $h \notin \mathcal{G}^2$ or $\mu_j^1 < 0$ for some $j \notin \mathcal{G}^1$ would imply the existence of an $\alpha > 0$ such that $y_h^0 + \alpha y_h^1 < 0$ or $\mu_j^0 + \alpha\mu_j^1 < 0$, thereby violating the constraints on y_h and μ_j in (8.8). Then $\sum_{j\notin\mathcal{G}^1} \mu_j^1 e(j) \geq 0$ and with (8.11) it follows that $\sum_{j\notin\mathcal{G}^2} y_j^1 a^j \geq 0$ and $y_j^1 \geq 0$ for all $j \notin \mathcal{G}^2$. Hence Assumption 8.1.2 implies that $y_j^1 = 0$ $(j \notin \mathcal{G}^2)$. Therefore, with (8.11), it follows that $\mu_j^1 = 0$ $(j \in \mathcal{G}^1)$.

But now we are left with the result that $\lambda_j^1 = 0$ $(j \in \mathcal{G}^1)$, $v_j^1 = 0$ $(j \in \mathcal{G}^2)$, $y_j^1 = 0$ $(j \notin \mathcal{G}^2)$, $\mu_j^1 = 0$ $(j \notin \mathcal{G}^1)$, $\theta^1 = 0$, and $\beta^1 = 0$. Hence no secondary ray can occur. □

The *SLCP*-algorithm introduced by Eaves computes a solution in each iteration of the sequence, thereby generating a sequence of prices $\{p^k\}_{k=0}^{\infty}$ possibly converging to an equilibrium price vector p^*. In each iteration of this *SLCP* a linear stationary point problem is solved. The algorithm follows a piecewise linear path of points starting in the price vector p^k either obtained as a solution to the previous iteration or, if no previous iteration exists, chosen arbitrarily from S_A^n. The path terminates within a finite number of pieces with a solution p^{k+1} of the linear stationary point problem problem. Each linear piece of this path is followed by generating a line segment of solutions to a system of $n + l + 2$ equations.

Chapter 9

An SLSPP-algorithm to compute an economic equilibrium

In this chapter we introduce a sequence of stationary point problems (*SLSPP*) to compute an equilibrium in an economy with linear production technologies introduced in Chapter 8 as an alternative to the *SLCP*-algorithms introduced in Mathiesen (1985b) and Eaves (1987) for this purpose. Mathiesen (1985b) shows that this equilibrium problem is equivalent to a nonlinear complementarity problem and he approximates this nonlinear complementarity problem by a sequence of linear complementarity problems. Each linear complementarity problem in this sequence is solved by the Lemke complementary pivoting algorithm. In this way a sequence of approximating solutions to the equilibrium problem is generated possibly converging to an equilibrium.

Since the Lemke complementary pivoting algorithm may fail to compute a solution to the linear complementarity problem obtained in an iteration of Mathiesen's *SLCP*, the generation of a sequence of approximating solutions cannot be assured. To overcome the possible divergence of the Lemke complementary pivoting algorithm as well as to improve the possibilities for a better understanding of the theoretical properties of the *SLCP*-algorithm, an alternative *SLCP*-algorithm was introduced in Eaves (1987). Eaves (1987) shows that the equilibrium problem in an economy with linear production technologies is equivalent to a stationary point problem on a polytope and he approximates this stationary point problem by a sequence of linear stationary

point problems. Each linear stationary point problem obtained in this sequence is shown to be solved thereby assuring the generation of a sequence of approximating solutions to the stationary point problem.

The algorithm introduced in this chapter approximates the stationary point problem by a sequence of linear stationary point problems. This results in a problem of a much lower dimension than the linear complementarity problem to be solved by Eaves in each iterate. The linear stationary point problem obtained in each iterate of the *SLSPP* is solved by an algorithm which is based on an algorithm introduced by Kamiya and Talman (1990) to compute a solution to linear stationary point problems on a polytope. Within a finite number of steps this algorithm solves the linear stationary point problem obtained in each iterate, thereby assuring the generation of a sequence of approximating solutions possibly converging to a solution of the equilibrium problem. This algorithm also avoids degeneracy problems that might occur in Eaves' *SLCP*.

This chapter is based on Kremers and Talman (1991) and is subdivided in the following way. In Section 1 we describe the sequence of linear stationary point problems while in Section 2 the steps of the algorithm to solve the linear stationary point problem in each iteration is given. Section 3 discusses convergency issues concerning the algorithm and shows that the method always succeeds to compute a solution to the linear stationary point problem obtained in each iteration of the *SLSPP*. In Section 4 we compare our *SLSPP*-algorithm with the *SLCP*-algorithms of Mathiesen and Eaves.

9.1 The sequence of linear stationary point problems

To compute an equilibrium in an exchange economy with $n+1$ commodities and l linear production technologies, let A be the $(n+1) \times l$ activity matrix and let the $(n+1)$-vector $z(p)$ be the total excess demand vector of the commodities at price vector $p \in \mathbf{R}_+^{n+1} \setminus \{0\}$. First we normalize the prices on the unit-simplex S^n. Secondly we only take into account the prices for which $p^\top A \leq 0$. The set of prices $p \in S^n$ such that $p^\top A \leq 0$ was denoted by S_A^n in the previous chapter and has the shape

of a polytope. This makes the problem of computing an equilibrium in an economy with linear production technologies equivalent to the problem of finding a price vector $p^* \in S_A^n$ and a vector of activity levels $y^* \geq 0$ such that $z(p^*) - Ay^* \leq 0$.

In Theorem 8.3.1 it was shown that the problem to compute an equilibrium in an economy with linear production technologies is equivalent to finding a stationary point of the excess demand function z on S_A^n. In order to solve the stationary point problem of z on S_A^n we propose to approximate this nonlinear problem by a sequence of linear stationary point problems. This *SLSPP* consists of a sequence of iterates where in each iterate the nonlinear *SPP* is linearized in a price vector either obtained as a solution to the previous iterate in the sequence or, if no previous iterate exists, chosen arbitrarily from S_A^n.[1]

Consider iteration k of the *SLSPP*. In this iteration the algorithm linearizes the *SPP* by replacing the excess demand function z with its first-order Taylor expansion in the price vector obtained as a solution to the previous iterate in case $k \geq 1$ and chosen arbitrarily from S_A^n in case $k = 0$. Denote this price vector by p^k. We allow p^k to lie on the boundary of S_A^n and we assume that the excess demand function z is differentiable in each point $p \in S^n$. Then the first-order Taylor expansion of the excess demand function z in p^k, denoted z^k, exists and is equal to

$$z^k(p) = z(p^k) + \nabla z(p^k)p \tag{9.1}$$

for all $p \in S_A^n$ where $\nabla z(p^k)$ denotes the Jacobian matrix of first-order derivatives of z to p in p^k. Recall that $\nabla z(p^k)p^k = 0$.

In iteration k the algorithm solves the linear stationary point problem of z^k on S_A^n, denoted by $LSPP_k$. Notice that a solution to $LSPP_k$ is generally not an equilibrium price vector in the original economy with linear production technologies.

If we are able to solve $LSPP_k$ for each $k = 0, 1, 2, \ldots$, then the *SLSPP*-algorithm generates a sequence of prices $\{p^k\}_{k=0}^{\infty}$ possibly converging to a solution to the *SPP* of z on S_A^n and therefore to an equilibrium price vector. In Section 2 of this chapter we introduce an algorithm to solve the $LSPP_k$ for every $k = 0, 1, 2, \ldots$.

[1] $S_A^n \neq \emptyset$ as follows from Assumption 8.1.2 and Minkowsky's Separating Hyperplane Theorem: $Ay \geq 0$, $y \geq 0$ does not have a solution $y > 0$ implies that there exists a vector $p > 0$ such that $p^T A \leq 0$.

9.2 The algorithm to solve $LSPP_k$

The algorithm we propose to solve the $LSPP_k$ in iteration k is a piecewise linear path-following algorithm starting in $p^k \in S_A^n$ and ending up with a stationary point of z^k on S_A^n. Each point p on the path can be seen as a stationary point of z^k on the set $S_A^n(\lambda) := (1-\lambda)\{p^k\} + \lambda S_A^n$ for some λ between zero and one. The parameter λ is a homotopy parameter running from zero to one. If $\lambda = 0$ then $S_A^n(0) = \{p^k\}$. So p^k is a stationary point of z^k on $S_A^n(0)$ and an end point of the path followed by the algorithm. When, while following the path, λ becomes one in a point $\overline{p} \in S_A^n$ then $\overline{p}$ is a stationary point of z^k on $S_A^n = S_A^n(1)$ and $p^{k+1} = \overline{p}$ is taken as the starting point for the next iteration. Figure 9.2.1 illustrates $S_A^n(\lambda)$ for several λ's when $n = 2$.

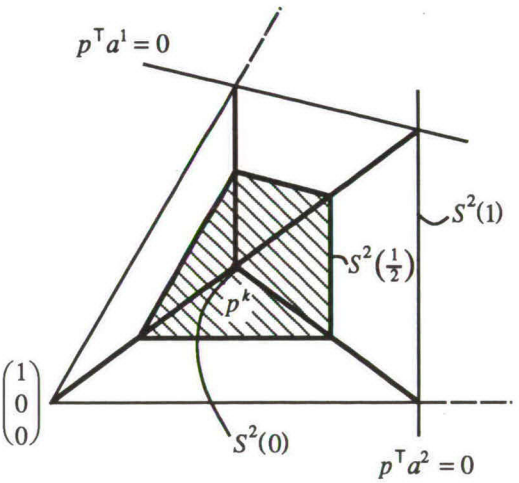


FIGURE 9.2.1: $S^2(\lambda)$ for $\lambda = 0, \frac{1}{2}, 1$.

Let p be an arbitrary point in $S_A^n(\lambda)$ for some given λ between zero and one. Then, by definition of $S_A^n(\lambda)$, $p = (1-\lambda)p^k + \lambda q$ for some $q \in S_A^n$. If q is such that it solves $\max\{\hat{p}^T z^k(p) \mid \hat{p} \in S_A^n\}$ then Lemma 9.2.1 shows that p is a stationary point of z^k on $S_A^n(\lambda)$.

Lemma 9.2.1 *For a point $p \in S_A^n$ it holds that $p = (1-\lambda)p^k + \lambda q$ for some λ, $0 \leq \lambda \leq 1$, with $q = \arg\max\{\hat{p}^\top z^k(p) \mid \hat{p} \in S_A^n\}$ if and only if p is a stationary point of z^k on $S_A^n(\lambda)$.*

Proof: Let $p = (1-\lambda)p^k + \lambda q$ for some λ, $0 \leq \lambda \leq 1$, such that $q = \arg\max\{\hat{p}^\top z^k(p) \mid \hat{p} \in S_A^n\}$. For every $\hat{p} \in S_A^n(\lambda)$ there exists a $\hat{q} \in S_A^n$ such that $\hat{p} = (1-\lambda)p^k + \lambda\hat{q}$ by definition of $S_A^n(\lambda)$. Consequently,

$$\begin{aligned}
\hat{p}^\top z^k(p) &= (1-\lambda)(p^k)^\top z^k(p) + \lambda\hat{q}^\top z^k(p) \\
&\leq (1-\lambda)(p^k)^\top z^k(p) + \lambda q^\top z^k(p) \\
&= ((1-\lambda)p^k + \lambda q)^\top z^k(p) \\
&= p^\top z^k(p).
\end{aligned}$$

Hence p is a stationary point of z^k on $S_A^n(\lambda)$. The converse follows from the discussions above. □

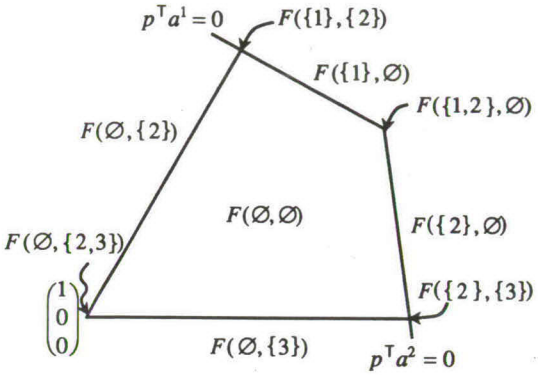
To characterize a stationary point of z^k on S_A^n and $S_A^n(\lambda)$, notice that a face of S_A^n is determined by

$$\begin{aligned}
\mathcal{F}(\mathcal{G}^1, \mathcal{G}^2) = \{p \in S_A^n \mid\ & p^\top a^j = 0 \text{ for all } j \in \mathcal{G}^1 \text{ and} \\
& p_j = 0 \text{ for all } j \in \mathcal{G}^2\}
\end{aligned}$$

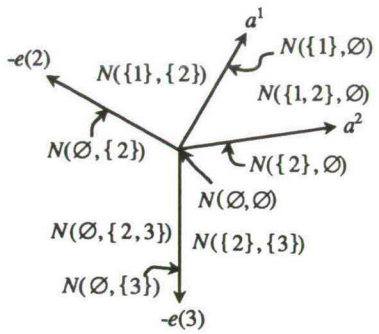
for certain sets $\mathcal{G}^1 \subseteq \mathcal{I}_l$ and $\mathcal{G}^2 \subset \mathcal{I}_{n+1}$. We assume without loss of generality that S_A^n is simple, so if $\mathcal{F}(\mathcal{G}^1, \mathcal{G}^2) \neq \emptyset$ then its dimension is equal to $n - |\mathcal{G}^1| - |\mathcal{G}^2|$. The normal cone of a nonempty set $\mathcal{F}(\mathcal{G}^1, \mathcal{G}^2)$, denoted $\mathcal{N}(\mathcal{G}^1, \mathcal{G}^2)$, is then defined by

$$\begin{aligned}
\mathcal{N}(\mathcal{G}^1, \mathcal{G}^2) = \{&\textstyle\sum_{j\in\mathcal{G}^1} y_j a^j - \sum_{j\in\mathcal{G}^2} \mu_j e(j) + \beta e \mid \\
& y_j \geq 0\ (j \in \mathcal{G}^1),\ \mu_j \geq 0\ (j \in \mathcal{G}^2), \text{ and } \beta \in \mathbf{R}\}.
\end{aligned}$$

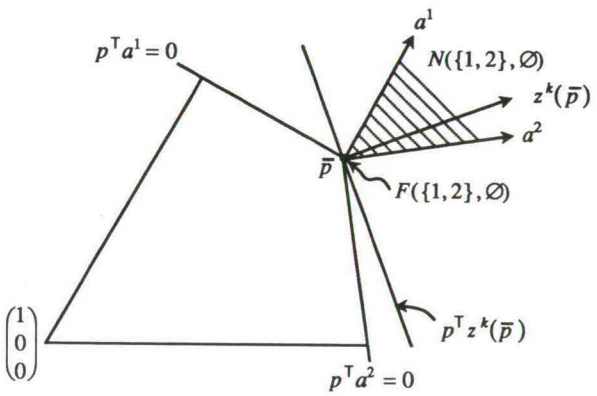
In Figure 9.2.2 we have illustrated these sets for the example of S_A^n in Figure 9.2.1, where the cones $\mathcal{N}(\mathcal{G}^1, \mathcal{G}^2)$ are drawn for β equal to zero. Theorem 9.2.1 states the relation between on the one hand $\mathcal{F}(\mathcal{G}^1, \mathcal{G}^2)$ and $\mathcal{N}(\mathcal{G}^1, \mathcal{G}^2)$, and on the other hand the stationary points of z^k on S_A^n, being as suggested in Figure 9.2.2.



(a) Subdivision of S_A^2 into sets $\mathcal{F}(\mathcal{G}^1,\mathcal{G}^2)$ for $\mathcal{G}^1 \subseteq \mathcal{I}_2$ and $\mathcal{G}^2 \subset \mathcal{I}_3$.



(b) The projected cones $\mathcal{N}(\mathcal{G}^1,\mathcal{G}^2)$ for $\mathcal{G}^1 \subseteq \mathcal{I}_2$ and $\mathcal{G}^2 \subset \mathcal{I}_3$.



(c) The point $\bar{p} \in \mathcal{F}(\{1,2\},\emptyset)$ maximizes $p^T z^k(\bar{p})$ over $p \in S_A^n$ if $z^k(\bar{p}) \in \mathcal{N}(\{1,2\},\emptyset)$.

FIGURE 9.2.2: The faces $\mathcal{F}(\mathcal{G}^1,\mathcal{G}^2)$ and their normal cones $\mathcal{N}(\mathcal{G}^1,\mathcal{G}^2)$ of S_A^2 for $\mathcal{G}^1 \subseteq \mathcal{I}_2$ and $\mathcal{G}^2 \subset \mathcal{I}_3$. In (c) Theorem 9.2.1 is illustrated at $\bar{p}$.

Theorem 9.2.1 *A point $\overline{p} \in S_A^n$ is a stationary point of z^k on S_A^n if and only if there exist sets $\mathcal{G}^1 \subseteq \mathcal{I}_l$ and $\mathcal{G}^2 \subset \mathcal{I}_{n+1}$ such that $\overline{p} \in \mathcal{F}(\mathcal{G}^1, \mathcal{G}^2)$ and $z^k(\overline{p}) \in \mathcal{N}(\mathcal{G}^1, \mathcal{G}^2)$.*

Proof: Suppose $\overline{p} \in S_A^n$ is a stationary point of z^k on S_A^n. Then $\overline{p}$ solves the linear programming problem, called the Primal, given by

$$\begin{aligned} &\max p^\top z^k(\overline{p}) \\ &\text{s.t.} \quad A^\top p \leq 0 \\ &\qquad e^\top p = 1 \\ &\qquad p \geq 0. \end{aligned}$$

The Duality Theorem of Linear Programming implies that this maximization problem is equivalent to solving the minimization problem, called the Dual, given by

$$\begin{aligned} &\min \ \beta \\ &\text{s.t.} \quad z^k(\overline{p}) = Ay + \beta e - \mu \\ &\qquad \mu \geq 0, \ y \geq 0, \ \beta \in \mathbf{R}. \end{aligned}$$

As $\overline{p}$ is a solution to the Primal there exists a unique solution $\overline{y}$, $\overline{\beta}$, and $\overline{\mu}$ to the maximization problem. Choose $\mathcal{G}^1$ to be the set $\{j \mid \overline{y}_j > 0\}$ and $\mathcal{G}^2$ to be the set $\{j \mid \overline{\mu}_j > 0\}$. Then the Dual implies that $z^k(\overline{p}) \in \mathcal{N}(\mathcal{G}^1, \mathcal{G}^2)$ and the complementarity property in linear programming gives $\overline{p} \in \mathcal{F}(\mathcal{G}^1, \mathcal{G}^2)$. □

This theorem says that, in order to solve $LSPP_k$, the algorithm could find a point $\overline{p} \in S_A^n$ such that $\overline{p} \in \mathcal{F}(\mathcal{G}^1, \mathcal{G}^2)$ and $z^k(\overline{p}) \in \mathcal{N}(\mathcal{G}^1, \mathcal{G}^2)$ for some subsets $\mathcal{G}^1 \subseteq \mathcal{I}_l$ and $\mathcal{G}^2 \subset \mathcal{I}_{n+1}$. Now let p and q be as in Lemma 9.2.1. Choose $\mathcal{G}^1 = \{j \in \mathcal{I}_l \mid q^\top a^j = 0\}$ and $\mathcal{G}^2 = \{j \in \mathcal{I}_{n+1} \mid q_j = 0\}$. Then $q \in \mathcal{F}(\mathcal{G}^1, \mathcal{G}^2)$ and, by construction of q, $z^k(p) \in \mathcal{N}(\mathcal{G}^1, \mathcal{G}^2)$. Therefore the algorithm we propose follows a path of points in S_A^n such that for every point p on the path there exist subsets $\mathcal{G}^1 \subseteq \mathcal{I}_l$ and $\mathcal{G}^2 \subset \mathcal{I}_{n+1}$ satisfying

1) $p = (1 - \lambda)p^k + \lambda q$ for some λ, $0 \leq \lambda \leq 1$, and some $q \in \mathcal{F}(\mathcal{G}^1, \mathcal{G}^2)$

2) $z^k(p) \in \mathcal{N}(\mathcal{G}^1, \mathcal{G}^2)$.

For an illustration we refer to Figure 9.2.3. Clearly p solves $LSPP_k$ when $\lambda = 1$ or $p^k \in \mathcal{F}(\mathcal{G}^1, \mathcal{G}^2)$, since in these cases also $p \in \mathcal{F}(\mathcal{G}^1, \mathcal{G}^2)$. Suppose $p^k \notin \mathcal{F}(\mathcal{G}^1, \mathcal{G}^2)$ and $\dim(\mathcal{F}(\mathcal{G}^1, \mathcal{G}^2)) = m$. We can represent each point $q \in \mathcal{F}(\mathcal{G}^1, \mathcal{G}^2)$ as an affine

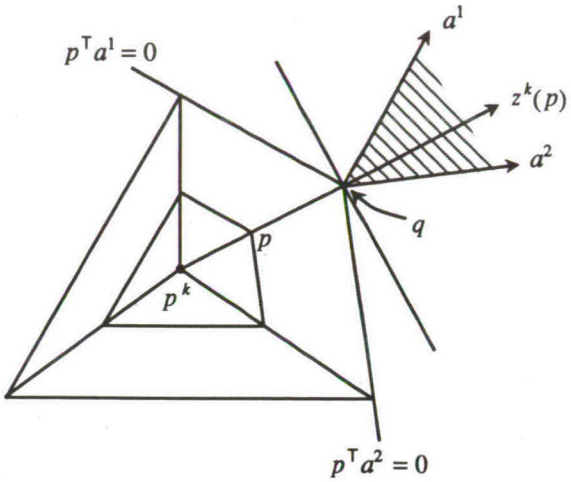


FIGURE 9.2.3: If the point p lies on the path generated by the algorithm then $p = (1-\lambda)p^k + \lambda q$ with $0 < \lambda < 1$ and $q \in \mathcal{F}(\{1,2\},\emptyset)$, while $z^k(p) \in \mathcal{N}(\{1,2\},\emptyset)$.

combination of $m+1$ affinely independent given points in the affine hull of $\mathcal{F}(\mathcal{G}^1,\mathcal{G}^2)$, aff$(\mathcal{F}(\mathcal{G}^1,\mathcal{G}^2))$. Let $w^0,\ldots,w^m$ represent these affinely independent points in aff$(\mathcal{F}(\mathcal{G}^1,\mathcal{G}^2))$. Then for every given $q \in \mathcal{F}(\mathcal{G}^1,\mathcal{G}^2)$ there exist numbers ν_j, $j = 0,1,\ldots,m$, such that $q = \sum_{j=0}^m \nu_j w^j$ and $\sum_{j=0}^m \nu_j = 1$. For some given λ, $0 \le \lambda \le 1$, let $\lambda_j = \lambda\nu_j$ for $j = 0,1,\ldots,m$. Then $\lambda q = \sum_{j=0}^m \lambda_j w^j$ and $\sum_{j=0}^m \lambda_j = \lambda$. Hence, the algorithm follows a path of points in S_A^n such that for every point p on the path there exist subsets $\mathcal{G}^1 \subseteq \mathcal{I}_l$ and $\mathcal{G}^2 \subset \mathcal{I}_{n+1}$ satisfying

1) $p = (1 - \sum_{j=0}^m \lambda_j)p^k + \sum_{j=0}^m \lambda_j w^j$, for some $\lambda_j \in \mathbf{R}$, $j \in \mathcal{I}_m \cup \{0\}$, and $0 \le \sum_{j=0}^m \lambda_j \le 1$, (9.2)

2) $z^k(p) \in \mathcal{N}(\mathcal{G}^1,\mathcal{G}^2)$,

where $w^0, w^1, \ldots, w^m$ are $m+1$ affinely independent points in aff$(\mathcal{F}(\mathcal{G}^1,\mathcal{G}^2))$.

Every point $p \in S_A^n$ satisfying (9.2) is a stationary point of z^k on $S_A^n(\lambda)$ with $\lambda = \sum_{j=0}^m \lambda_j$. If one combines (9.2) with $z^k(p) = z(p^k) + \nabla z(p^k)p$, then p is a stationary

point of z^k on $S_A^n(\lambda)$ for some λ, $0 \leq \lambda \leq 1$, if and only if for some $\mathcal{G}^1 \subseteq \mathcal{I}_l$ and $\mathcal{G}^2 \subset \mathcal{I}_{n+1}$ the system of linear equations

$$\sum_{j=0}^{m} \lambda_j \nabla z(p^k) w^j - \sum_{j \in \mathcal{G}^1} y_j a^j + \sum_{j \in \mathcal{G}^2} \mu_j e(j) - \beta e = -z(p^k) \tag{9.3}$$

has a solution $\lambda_j \in \mathbf{R}$ $(j \in \mathcal{I}_m \cup \{0\})$, $y_j \geq 0$ $(j \in \mathcal{G}^1)$, $\mu_j \geq 0$ $(j \in \mathcal{G}^2)$, $\beta \in \mathbf{R}$, such that $0 < \sum_{j=0}^{m} \lambda_j \leq 1$, $p = (1-\lambda)p^k + \lambda q$, $q^\top a^j \leq 0$ for $j \notin \mathcal{G}^1$, and $q_j \geq 0$ for $j \notin \mathcal{G}^2$, where $q = \sum_{j=0}^{m} \lambda_j w^j / \lambda$ and $\lambda = \sum_{j=0}^{m} \lambda_j$. This system contains $n+1$ equations with $n+2$ unknowns leaving us with one degree of freedom. Therefore, assuming nondegeneracy, the set of solutions to system (9.3) if nonempty represents a line segment of solutions to the *LSPP* of z^k on $S_A^n(\lambda)$, $0 \leq \lambda \leq 1$.

As will be shown in Section 3 and by assuming nondegeneracy the line segment of solutions obtained from (9.3) has two end points. This line segment will be followed by making a linear programming pivot step in (9.3) in one end point with one of the variables λ_m, y_j for some $j \in \mathcal{G}^1$, or μ_j for some $j \in \mathcal{G}^2$. The other end point of this line segment is a point $\overline{p}$ in S_A^n where either $\sum_{j=0}^{m} \lambda_j = 1$, or one of the restricted variables in (9.3) is equal to zero, or one of the constraints on q is binding.

Case 1: $\sum_{j=0}^{m} \lambda_j$ becomes equal to 1. Then $\overline{p} = \sum_{j=0}^{m} \lambda_j w^j \in \mathcal{F}(\mathcal{G}^1, \mathcal{G}^2)$ while $z^k(\overline{p}) \in \mathcal{N}(\mathcal{G}^1, \mathcal{G}^2)$. Theorem 9.2.1 implies that $\overline{p}$ is a stationary point of z^k on S_A^n and $p^{k+1} = \overline{p}$ is taken as the starting point to the next iteration.

Case 2: y_t becomes zero for some $t \in \mathcal{G}^1$. Then, at $\overline{p}$, it holds that

$$z^k(\overline{p}) = \sum_{j \in \mathcal{G}^1 \setminus \{t\}} y_j a^j - \sum_{j \in \mathcal{G}^2} \mu_j e(j) + \beta e.$$

Hence $z^k(\overline{p}) \in \mathcal{N}(\mathcal{G}^1 \setminus \{t\}, \mathcal{G}^2)$. If $p^k \in \mathcal{F}(\mathcal{G}^1 \setminus \{t\}, \mathcal{G}^2)$, i.e. $(p^k)^\top a^j = 0$ for all $j \in \mathcal{G}^1 \setminus \{t\}$ and $p_j^k = 0$ for all $j \in \mathcal{G}^2$, then $\overline{p} \in \mathcal{F}(\mathcal{G}^1 \setminus \{t\}, \mathcal{G}^2)$ and Theorem 9.2.1 implies that $\overline{p}$ is a stationary point of z^k on S_A^n. Subsequently $p^{k+1} = \overline{p}$ is taken as the starting point to the next iteration. If $p^k \notin \mathcal{F}(\mathcal{G}^1 \setminus \{t\}, \mathcal{G}^2)$, then the algorithm maintains the validity of the conditions in (9.2) by generating prices $p = (1-\lambda)p^k + \lambda q$ such that $0 \leq \lambda \leq 1$, $z^k(p) \in \mathcal{N}(\mathcal{G}^1 \setminus \{t\}, \mathcal{G}^2)$, and the vector q is an affine combination in $\mathcal{F}(\mathcal{G}^1 \setminus \{t\}, \mathcal{G}^2)$ of $m+2$ affinely independent points $\hat{w}^0, \ldots, \hat{w}^{m+1}$ in $\mathrm{aff}(\mathcal{F}(\mathcal{G}^1 \setminus \{t\}, \mathcal{G}^2))$. Since the points $w^0, \ldots, w^m$ are already affinely independent points in $\mathrm{aff}(\mathcal{F}(\mathcal{G}^1, \mathcal{G}^2))$ and $\mathrm{aff}(\mathcal{F}(\mathcal{G}^1 \setminus \{t\}, \mathcal{G}^2))$ is a subset of $\mathrm{aff}(\mathcal{F}(\mathcal{G}^1, \mathcal{G}^2))$, one can take $\hat{w}^j = w^j$ for $j = 0, 1, \ldots, m$. Then $\hat{w}^{m+1}$ must be such that $\hat{w}^{m+1} \in \mathrm{aff}(\mathcal{F}(\mathcal{G}^1 \setminus \{t\}, \mathcal{G}^2))$ and

$\hat{w}^{m+1} \notin \text{aff}(\mathcal{F}(\mathcal{G}^1, \mathcal{G}^2))$. By definition this implies that $\hat{w}^{m+1}$ has to fulfil the conditions

$$\begin{aligned} (a^j)^\top \hat{w}^{m+1} &= 0 \text{ for all } j \in \mathcal{G}^1 \setminus \{t\} \\ (a^t)^\top \hat{w}^{m+1} &\neq 0 \\ e^\top \hat{w}^{m+1} &= 1 \\ \hat{w}_j^{m+1} &= 0 \text{ for all } j \in \mathcal{G}^2. \end{aligned}$$

For example we can take $(a^t)^\top \hat{w}^{m+1} = -1$.

Let C be the matrix consisting of the basic column vectors $-e$, $-a^j$ $(j \in \mathcal{G}^1)$, and $e(j)$ $(j \in \mathcal{G}^2)$ of system (9.3). Then the conditions on $\hat{w}^{m+1}$ reduce to

$$C^\top \hat{w}^{m+1} = \hat{e}(h_1) - \hat{e}(h_2),$$

where h_2 denotes the index of the column $-e$ in C, h_1 denotes the index of the column $-a^t$ in C, and $\hat{e}(j)$ is the j-th unit vector having the same length as the number of columns in the matrix C. Let D denote the matrix containing the remaining basic vectors of (9.3) as its columns. Then $\hat{w}^{m+1}$ can be determined from the system

$$\begin{pmatrix} C^\top \\ D^\top \end{pmatrix} \hat{w}^{m+1} = \begin{pmatrix} \hat{e}(h_1) - \hat{e}(h_2) \\ d \end{pmatrix} \tag{9.4}$$

for some arbitrarily chosen vector d of appropriate length. Without loss of generality we take $d = 0$. Then $\hat{w}^{m+1}$ is equal to

$$\hat{w}^{m+1} = (B^\top)^{-1} e(h_1) - (B^\top)^{-1} e(h_2) \tag{9.5}$$

where $B = (C \ D)$, $e(h_1) = (\hat{e}(h_1)^\top, \ 0)^\top$ and $e(h_2) = (\hat{e}(h_2)^\top, \ 0)^\top$. Notice that the inverse of the matrix $B^\top$ is the transpose of the basis inverse obtained from the pivoting tableau corresponding to system (9.3). To obtain $\hat{w}^{m+1}$ we must therefore subtract the row of the basis inverse corresponding to a^t from the row of this matrix corresponding to e. The algorithm proceeds by increasing m with one and pivoting the column $\nabla z(p^k)\hat{w}^m$ into the appropriate pivot system (9.3) thereby raising the variable λ_m from zero.

Case 3: μ_t becomes zero for some $t \in \mathcal{G}^2$ in a point $\overline{p} \in S_A^n$. This case is similar to the previous one. In $\overline{p}$ it holds that

$$z^k(\overline{p}) = \sum_{j \in \mathcal{G}^1} y_j a^j - \sum_{j \in \mathcal{G}^2 \setminus \{t\}} \mu_j e(j) + \beta e.$$

Hence $z^k(\overline{p}) \in \mathcal{N}(\mathcal{G}^1, \mathcal{G}^2 \setminus \{t\})$. If $p^k \in \mathcal{F}(\mathcal{G}^1, \mathcal{G}^2 \setminus \{t\})$, i.e. $(p^k)^\top a^j = 0$ for all $j \in \mathcal{G}^1$ and $p_j^k = 0$ for all $j \in \mathcal{G}^2 \setminus \{t\}$, then $\overline{p} \in \mathcal{F}(\mathcal{G}^1, \mathcal{G}^2 \setminus \{t\})$ and Theorem 9.2.1 implies that $\overline{p}$ is a stationary point of z^k on S_A^n. Subsequently $p^{k+1} = \overline{p}$ is taken as the starting point to the next iteration. If $p^k \notin \mathcal{F}(\mathcal{G}^1, \mathcal{G}^2 \setminus \{t\})$, then the algorithm maintains the validity of the conditions in (9.2) by generating prices $p = (1-\lambda)p^k + \lambda q$ such that $0 \leq \lambda \leq 1$, $z^k(p) \in \mathcal{N}(\mathcal{G}^1, \mathcal{G}^2 \setminus \{t\})$, and the vector q is an affine combination in $\mathcal{F}(\mathcal{G}^1, \mathcal{G}^2 \setminus \{t\})$ of $m+2$ affinely independent given points $\hat{w}^0, \ldots, \hat{w}^{m+1}$ in $\text{aff}(\mathcal{F}(\mathcal{G}^1, \mathcal{G}^2 \setminus \{t\}))$. Similar to *Case 2* we take $\hat{w}^j = w^j$, $j = 0, 1, \ldots, m$, and $\hat{w}^{m+1} = (B^\top)^{-1}e(h_1) - (B^\top)^{-1}e(h_2)$ where $(B^\top)^{-1}$ and h_1 are defined as in *Case 2* while h_2 corresponds to the index of the column $e(t)$. The algorithm proceeds by increasing m with one and pivoting the column $\nabla z(p^k)\hat{w}^m$ into the appropriate pivot system thereby raising the variable λ_m from zero.

Case 4: For some $t \notin \mathcal{G}^1$ it holds at $\overline{p}$ that

$$\sum_{j=0}^{m} \lambda_j (a^t)^\top w^j = 0.$$

Then $(a^t)^\top q^1 = 0$ where $q^1 = \sum_{j=0}^m \lambda_j w^j / \lambda$ is such that $\overline{p} = (1-\lambda)p^k + \lambda q^1$. Hence $q^1 \in \mathcal{F}(\mathcal{G}^1 \cup \{t\}, \mathcal{G}^2) \subset \text{bd}(\mathcal{F}(\mathcal{G}^1, \mathcal{G}^2))$. The algorithm maintains the validity of (9.2) by generating prices $p = (1-\lambda)p^k + \lambda q$ such that $0 \leq \lambda \leq 1$, $q \in \mathcal{F}(\mathcal{G}^1 \cup \{t\}, \mathcal{G}^2)$, and $z^k(p) \in \mathcal{N}(\mathcal{G}^1 \cup \{t\}, \mathcal{G}^2)$. Since $\dim(\mathcal{F}(\mathcal{G}^1 \cup \{t\}, \mathcal{G}^2)) = m-1$ we should determine m affinely independent points in $\text{aff}(\mathcal{F}(\mathcal{G}^1 \cup \{t\}, \mathcal{G}^2))$. These points, say $\hat{w}^0, \ldots, \hat{w}^{m-1}$, can be obtained by a parallel movement of the points $w^0, \ldots, w^m$ onto $\text{aff}(\mathcal{F}(\mathcal{G}^1 \cup \{t\}, \mathcal{G}^2))$ and deleting one of them.

Let $r = (1 - \sum_{j=0}^m \lambda_j^0)p^k + \sum_{j=0}^m \lambda_j^0 q^0$ be the previous end point obtained by the algorithm, where $q^0 \in \mathcal{F}(\mathcal{G}^1, \mathcal{G}^2)$ is such that $q^0 = \sum_{j=0}^m ((\lambda_j^0/\lambda^0)w^j)$ with $0 < \lambda_0 = \sum_{j=0}^m \lambda_j^0 \leq 1$ and $\lambda_j^0 \in \mathbf{R}$ obtained from the solution to (9.3) in r. The algorithm moves all points w^j, $j = 0, 1, \ldots, m$, parallel to $q^1 - q^0$ onto $\text{aff}(\mathcal{F}(\mathcal{G}^1 \cup \{t\}, \mathcal{G}^2))$. This parallel movement of w^j, $j = 0, 1, \ldots, m$, results in $m+1$ points $w^j + \delta^j(q^1 - q^0) \in \text{aff}(\mathcal{F}(\mathcal{G}^1 \cup \{t\}, \mathcal{G}^2))$ where $\delta^j = (a^t)^\top w^j / (a^t)^\top q^0$, $j = 0, 1, \ldots, m$. The appendix to this chapter shows that if we delete a point with index $g \in \{0, \ldots, m\}$ for which $\lambda_g \neq \lambda_g^0$ then the remaining points constitute m affinely independent points in $\text{aff}(\mathcal{F}(\mathcal{G}^1 \cup \{t\}, \mathcal{G}^2))$.

Take $\hat{w}^j = w^j + \delta^j(q^1 - q^0)$, $j = 0, 1, \ldots, g-1$, and $\hat{w}^{j-1} = w^j + \delta^j(q^1 - q^0)$,

$j = g+1,\ldots,m$. Then, in the current basis inverse we replace $\nabla z(p^k)w^j$ by the column $\nabla z(p^k)\hat{w}^j$, $j = 1,\ldots,g-1$, $\nabla z(p^k)w^j$ by $\nabla z(p^k)\hat{w}^{j-1}$ for $j = g+1,\ldots,m$, and $\nabla z(p^k)w^g$ by the vector with which the pivoting step was made. Notice that if the pivoting step was made with the vector $\nabla z(p^k)w^m$ then $\hat{w}^0,\ldots,\hat{w}^{m-1}$ must be affinely independent since $\lambda_m > 0 = \lambda_m^0$. In that case we can take $g = m$ and $\nabla z(p^k)w^m$ does not enter the basis inverse. The algorithm proceeds by decreasing m with one and pivoting the column $-a^t$ into the new pivot system thereby raising the variable y_t from zero .

Case 5: For some $t \notin \mathcal{G}^2$ it holds that

$$\sum_{j=0}^{m} \lambda_j w_t^j = 0.$$

This case is similar to the previous one. Then $q_t^1 = 0$ where $q^1 = \sum_{j=0}^m \lambda_j w^j/\lambda$ is such that $\overline{p} = (1-\lambda)p^k + \lambda q^1$ and $q^1 \in \mathcal{F}(\mathcal{G}^1,\mathcal{G}^2 \cup \{t\}) \subset \mathrm{bd}(\mathcal{F}(\mathcal{G}^1,\mathcal{G}^2))$. The algorithm maintains the validity of (9.2) by generating prices $p = (1-\lambda)p^k + \lambda q$ such that $0 \leq \lambda \leq 1$, $q \in \mathcal{F}(\mathcal{G}^1,\mathcal{G}^2 \cup \{t\})$, and $z^k(p) \in \mathcal{N}(\mathcal{G}^1,\mathcal{G}^2 \cup \{t\})$. Similar to *Case 4* we determine m affinely independent points in $\mathrm{aff}(\mathcal{F}(\mathcal{G}^1,\mathcal{G}^2 \cup \{t\}))$ from $w^j + \delta^j(q^1 - q^0)$ for $j = 0,1,\ldots,m$ where $\delta^j = w_t^j/q_t^0$ and q^0 as defined in *Case 4*, and adapt the current basis inverse. The algorithm proceeds by decreasing m with one and pivoting the column $e(t)$ into the new pivot system thereby raising the variable μ_t from zero.

The Cases 1 to 5 describe the performance of the algorithm at the end points of all possible line segments generated by the algorithm except at p^k where the algorithm is initiated. To show that p^k is an end point of a line segment of stationary points of z^k on $S_A^n(\lambda)$ for λ, $0 \leq \lambda \leq 1$, we have to find sets $\mathcal{G}^1 \subseteq \mathcal{I}_l$ and $\mathcal{G}^2 \subset \mathcal{I}_{n+1}$ such that (9.2) is satisfied in p^k. This means that according to Lemma 9.2.1 this line segment contains points $p \in S_A^n$ such that $p = (1-\lambda)p^k + \lambda q$ for some λ, $0 \leq \lambda \leq 1$, and some $q \in \mathcal{F}(\mathcal{G}^1,\mathcal{G}^2)$ maximizing $p^\top z^k(p^k)$ subject to $p \in S_A^n$. Clearly q follows from solving either

the Primal			or	*the Dual*		
$\max$	$p^\top z(p^k)$			$\min$	β	
s.t.	$A^\top p$	≤ 0		s.t.	$Ay - \mu + \beta e = z(p^k)$	
	$-p$	≤ 0			$y \geq 0,\ \mu \geq 0.$	
	$e^\top p$	$= 1$				

Let $\mathcal{G}_0^1 = \{j \mid y_j > 0\}$ and $\mathcal{G}_0^2 = \{j \mid \mu_j > 0\}$, after solving the Dual. Assuming nondegeneracy the sets $\mathcal{G}_0^1$ and $\mathcal{G}_0^2$ define a face $\mathcal{F}(\mathcal{G}_0^1, \mathcal{G}_0^2)$ being a vertex of S_A^n, say w^0. In case $w^0 = p^k$ then p^k is an equilibrium price vector for the original economy. Otherwise p^k is the end point of a line segment of points $p \in S_A^n$ such that

1) $p = (1 - \lambda_0)p^k + \lambda_0 w^0$, for some λ_0, $0 \leq \lambda_0 \leq 1$

2) $z^k(p) \in \mathcal{N}(\mathcal{G}_0^1, \mathcal{G}_0^2)$.

Hence p^k fulfils the conditions in (9.2).

Notice that combining these conditions leads to the system

$$\lambda_0 \nabla z(p^k)w^0 - \sum_{j \in \mathcal{G}_0^1} y_j a^j + \sum_{j \in \mathcal{G}_0^2} \mu_j e(j) - \beta e = -z(p^k) \tag{9.6}$$

having a solution $0 \leq \lambda_0 \leq 1$, $y_j \geq 0$ $(j \in \mathcal{G}_0^1)$, $\mu_j \geq 0$ $(j \in \mathcal{G}_0^2)$, and $\beta \in \mathbf{R}$. In $p = p^k$, system (9.6) has as unique solution $\lambda_0 = 0$ and $y_j > 0$ $(j \in \mathcal{G}_0^1)$, $\mu_j > 0$ $(j \in \mathcal{G}_0^2)$, $\beta > 0$, obtained from solving the Dual.

The line segment of solutions to (9.6) is followed from $p = p^k$ by making a linear programming pivot step with the column vector $\nabla z(p^k)w^0$ in (9.6) thereby raising λ_0 from zero. While raising λ_0 from zero the algorithm will encounter another end point of the line segment when either y_j becomes zero for some $j \in \mathcal{G}_0^1$, or μ_j becomes zero for some $j \in \mathcal{G}_0^2$, or λ_0 becomes one. The performance of the algorithm in this end point is described in the Cases 1 to 3 above. This concludes the description of the steps of the algorithm.

9.3 Convergency issues

Starting in the point p^k obtained from the previous iteration of the *SLSPP* or, in case no previous iteration exists, chosen arbitrarily from S_A^n, the algorithm follows in iteration k a path of points either ending up with a solution to the $LSPP_k$ or ending up in a ray. In the previous section we described the performance of the algorithm in all possible end points of the line segments of the path followed by the algorithm. During this description we encountered cases in which the algorithm ends up with a solution to the $LSPP_k$. These cases are summarized in Lemma 9.3.1.

Lemma 9.3.1 *Let $\bar{p} = (1-\lambda)p^k + \lambda\bar{q}$ be an end point of a line segment of points generated by the algorithm with $\bar{q} \in \mathcal{F}(\mathcal{G}^1, \mathcal{G}^2)$ and $z^k(\bar{p}) \in \mathcal{N}(\mathcal{G}^1, \mathcal{G}^2)$ for some $\mathcal{G}^1 \subseteq \mathcal{I}_l$ and $\mathcal{G}^2 \subset \mathcal{I}_{n+1}$. Then $\bar{p}$ is a solution to the $LSPP_k$ if one of the following cases for the corresponding solution of (9.3) holds:*

i) $\lambda = 1$;

ii) $y_t = 0$ *for some* $t \in \mathcal{G}^1$, $(p^k)^\top a^j = 0$ *for all* $j \in \mathcal{G}^1 \setminus \{t\}$, *and* $p_j^k = 0$ *for* $j \in \mathcal{G}^2$;

iii) $\mu_t = 0$ *for some* $t \in \mathcal{G}^2$, $(p^k)^\top a^j = 0$ *for all* $j \in \mathcal{G}^1$, *and* $p_j^k = 0$ *for* $j \in \mathcal{G}^2 \setminus \{t\}$.

The algorithm does not find a solution to the $LSPP_k$ if it ends up in a ray of solutions to system (9.3) for some $\mathcal{G}^1 \subseteq \mathcal{I}_l$ and $\mathcal{G}^2 \subset \mathcal{I}_{n+1}$. This means that the pivot variable can be raised towards infinity without violating any of the constraints on the variables in (9.3). Theorem 9.3.1 shows that the algorithm cannot end up in such a ray.

Theorem 9.3.1 *For all possible sets $\mathcal{G}^1 \subseteq \mathcal{I}_l$ and $\mathcal{G}^2 \subset \mathcal{I}_{n+1}$ and all $\hat{p} \in S_A^n$ the system of equations*

$$\sum_{j=0}^{m} \lambda_j \nabla z(\hat{p}) w^j - \sum_{j \in \mathcal{G}^1} y_j a^j + \sum_{j \in \mathcal{G}^2} \mu_j e(j) - \beta e = -z(\hat{p}) \tag{9.7}$$

does not contain a ray of solutions satisfying $\lambda_j \in \mathbf{R}$ $(j \in \mathcal{I}_m \cup \{0\})$, $y_j \geq 0$ $(j \in \mathcal{G}^1)$, $\mu_j \geq 0$ $(j \in \mathcal{G}^2)$, $\beta \in \mathbf{R}$, $0 < \lambda = \sum_{j=0}^{m} \lambda_j \leq 1$, $q^\top a^j \leq 0$ for $j \notin \mathcal{G}^1$, $q_j \geq 0$, for $j \notin \mathcal{G}^2$ where $q = \sum_{j=0}^{m} \lambda_j \lambda^{-1} w^j$.

Proof: Suppose (9.7) contains a ray of admissable solutions $\lambda_j^0 + \alpha\lambda_j^1$ $(j \in \mathcal{I}_m \cup \{0\})$, $y_j^0 + \alpha y_j^1$ $(j \in \mathcal{G}^1)$, $\mu_j^0 + \alpha\mu_j^1$ $(j \in \mathcal{G}^2)$, $\beta^0 + \alpha\beta^1$ for all $\alpha \geq 0$. Then λ_j^0 $(\mathcal{I}_m \cup \{0\})$, y_j^0 $(j \in \mathcal{G}^1)$, μ_j^0 $(j \in \mathcal{G}^2)$, β^0 is an admissable solution to (9.7) and λ_j^1 $(\mathcal{I}_m \cup \{0\})$, y_j^1 $(j \in \mathcal{G}^1)$, μ_j^1 $(j \in \mathcal{G}^2)$, β^1 is a solution to the homogeneous system of equations

$$\sum_{j=0}^{m} \lambda_j \nabla z(\hat{p}) w^j - \sum_{j \in \mathcal{G}^1} y_j a^j + \sum_{j \in \mathcal{G}^2} \mu_j e(j) - \beta e = 0. \tag{9.8}$$

It is obvious that $\sum_{j=0}^{m} \lambda_j^1 = 0$ as $\sum_{j=0}^{m} \lambda_j^1 \neq 0$ would imply that there exists some $\alpha > 0$ such that $\sum_{j=0}^{m} \lambda_j^0 + \alpha(\sum_{j=0}^{m} \lambda_j^1) > 1$ or $\sum_{j=0}^{m} \lambda_j^0 + \alpha(\sum_{j=0}^{m} \lambda_j^1) < 0$. Let $\lambda^0 = \sum_{j=0}^{m} \lambda_j^0$, so $\lambda^0 > 0$. Let $q(\alpha) = \sum_{j=0}^{m} (\lambda_j^0 + \alpha\lambda_j^1) w^j / \lambda^0$ then $q(\alpha)$ must be in $\mathcal{F}(\mathcal{G}^1, \mathcal{G}^2)$

for every $\alpha > 0$. Since $\mathcal{F}(\mathcal{G}^1, \mathcal{G}^2)$ is compact this implies that $q(\alpha) = q(0)$ for all $\alpha > 0$, i.e. $\sum_{j=0}^{m} \lambda_j^1 w^j = 0$. By affine independence of w^j ($j \in \mathcal{I}_m \cup \{0\}$), this implies together with $\sum_{j=0}^{m} \lambda_j^1 = 0$ that $\lambda_j^1 = 0$ for all $j \in \mathcal{I}_m \cup \{0\}$. Premultiplying (9.8) with $q^0 = \sum_{j=0}^{m} \lambda_j^0 (\lambda^0)^{-1} w^j$ gives $\beta^1 = 0$ since $(q^0)^T a^j = 0$ for $j \in \mathcal{G}^1$ and $q_j^0 = 0$ for $j \in \mathcal{G}^2$. Then it follows that

$$\sum_{j \in \mathcal{G}^2} \mu_j^1 e(j) = \sum_{j \in \mathcal{G}^1} y_j^1 a^j. \tag{9.9}$$

Notice that $y_j^1 \geq 0$ ($j \in \mathcal{G}^1$) as well as $\mu_j^1 \geq 0$ ($j \in \mathcal{G}^2$). This follows from $y_j^0 + \alpha y_j^1$ ($j \in \mathcal{G}^1$) and $\mu_j^0 + \alpha \mu_j^1$ ($j \in \mathcal{G}^2$) being a solution to (9.7) for all $\alpha \geq 0$. Indeed, as $y_h^1 < 0$ for some $h \in \mathcal{G}^1$ or $\mu_j^1 < 0$ for some $j \in \mathcal{G}^2$ would imply the existence of an $\alpha > 0$ such that $y_h^0 + \alpha y_h^1 < 0$ or $\mu_j^0 + \alpha \mu_j^1 < 0$, thereby violating the constraints on y_h and μ_j in (9.7). Then $\sum_{j \in \mathcal{G}^2} \mu_j^1 e(j) \geq 0$ and with (9.9) it follows that $\sum_{j \in \mathcal{G}^1} y_j^1 a^j \geq 0$ and $y_j^1 \geq 0$ for all $j \in \mathcal{G}^1$. Hence Assumption 8.1.2 implies that $y_j^1 = 0$ ($j \in \mathcal{G}^1$). Therefore, with (9.9), it follows that $\mu_j^1 = 0$ ($j \in \mathcal{G}^2$).

But now we are left with the result that $\lambda_j^1 = 0$ ($j \in \mathcal{I}_m \cup \{0\}$), $y_j^1 = 0$ ($j \in \mathcal{G}^1$), $\mu_j^1 = 0$ ($j \in \mathcal{G}^2$), and $\beta^1 = 0$. Hence no ray of solutions to (9.3) can occur for any $\mathcal{G}^1 \subseteq \mathcal{I}_l$ and $\mathcal{G}^2 \subset \mathcal{I}_{n+1}$. □

Theorem 9.3.1 guarantees that each line segment on the path contains exactly two end points. Assuming nondegeneracy, at an end point of a line segment just one of the five cases described in Section 2 can occur. Therefore the starting point is the end point of a unique line segment whereas each other end point of a line segment is either an end point of a uniquely determined other line segment or a solution to $LSPP_k$.

Every line segment on the path in S_A^n is determined by the line segment of solutions to (9.3) for some unique $\mathcal{G}^1 \subseteq \mathcal{I}_l$ and $\mathcal{G}^2 \subset \mathcal{I}_{n+1}$. As $\mathcal{G}^1$ and $\mathcal{G}^2$ are both subsets of finite sets the total number of line segments is finite. Hence starting in $p = p^k$ the algorithm generates a finite sequence of different line segments. Therefore the algorithm terminates within a finite number of steps with a solution to $LSPP_k$.

9.4 Conclusions

In this chapter we have introduced an *SLSPP*-algorithm to compute an equilibrium in an economy with linear production technologies. This *SLSPP*-algorithm consists of iterations in which the stationary point problem determining an equilibrium in an economy with linear production technologies is linearized by taking the first-order Taylor expansion of the excess demand function z in the price vector either obtained as a solution from the previous iteration or, if no previous iteration exists, in an arbitrarily chosen price vector from S_A^n. In each iteration this results in a linear stationary point problem which we solve by the algorithm introduced in Section 2. In Theorem 9.3.1 we proved that this algorithm finds a solution for all possible starting points in S_A^n. Therefore the *SLSPP*-algorithm we introduced in this chapter certainly generates a sequence of prices $\{p^k\}_{k=0}^{\infty}$, possibly converging towards an equilibrium in an economy with linear production technologies.

Contrary to the existence of convergence to an approximating price vector p^{k+1} in each iteration k of the algorithm we are not able to show global convergence. In Mathiesen (1987) some empirical results are given suggesting global convergence for his *SLCP*-algorithm. However, Mathiesen (1985c) also gave some examples, namely Scarf's unstable equilibria (see Scarf (1960) for details), where his *SLCP*-algorithm failed to converge. Mathiesen's *SLCP*-algorithm does not even need to converge in any iteration, thereby possibly failing to generate a sequence of prices $\{p^k\}_{k=0}^{\infty}$. Furthermore Mathiesen has to choose among different *LCP*s to solve (dependent on the numeraire choice) in each iteration.

Our algorithm can be seen as an improvement of the algorithm introduced in Eaves (1987). Eaves normalizes prices on the unit simplex in each iteration instead of taking some commodity as a numeraire. Then Eaves introduces an algorithm which generates a path in S_A^n by incorporating the "no-profit"-conditions in the pivot system. This results in a pivot system consisting of $n + l + 2$ equations in each iteration contrary to our pivot system which just consists of the "market-clearance"-conditions and hence results in a system of only $n + 1$ equations. Therefore our algorithm processes the same information more efficiently than Eaves' algorithm. This can easily be seen by taking an economy where the number of activities largely uotnumbers the number of commodities.

The possibility that the algorithm has to start at the beginning of some iteration on the boundary with the price of some commodities equal to zero may cause problems in Mathiesen's *SLCP*, Eaves' *SLCP* as well as our *SLSPP* because in general $z_j(p) \to \infty$ when p_j converges to zero. Mathiesen and Eaves suggest to perturbe such a p^k by taking the modified price vector $p^{k-1} + t(p^k - p^{k-1})$, for some t, $0 < t < 1$, instead of p^k as a starting point for iteration k. This perturbation of the starting point can cause the algorithm to zigzag along the boundary as some of Mathiesen's numerical examples showed. This might also occur in our algorithm.

As Eaves algorithm as well as our algorithm generate only prices in the subset S_A^n of S^n it may also occur that in some iteration the starting point lies in the boundary of S_A^n with some binding "no-profit"-conditions. This causes no problems for our algorithm. The performance remains exactly the same as described in Section 2. In Eaves' algorithm the starting point has to be perturbed to prevent degeneracy in the starting point. Because this case may often occur when applying the algorithm this feature of our algorithm can be regarded as an improvement of Eaves' *SLCP*.

In Mathiesen (1985b) the Lemke algorithm (see Lemke (1965)) was applied to solve the *LCP* in each of the iterations of the *SLCP*. The main drawback of applying Lemke's algorithm is that it lacks the possibility to start in an arbitrary starting point, causing a loss of information when proceeding from one iteration of the *SLCP* to the next one. Furthermore Mathiesen (1985b) had to restate the original equilibrium problem to a form suitable to apply the Lemke algorithm. This last feature as well as the fixed starting point imposed by applying the Lemke algorithm contributed to a large extent in the problems encountered when using the *SLCP* of Mathiesen.

In order to get rid of the problems encountered with Mathiesen (1985b) and Eaves (1987) we showed that the original equilibrium problem in an economy with linear production technologies is equivalent to the stationary point problem of z on the subset S_A^n of S^n instead of rewriting this problem to a form fit for applying an already existing algorithm. This *SPP* was approximated by a sequence of linear stationary point problems. To solve the linear stationary point problem in each iteration of the sequence we introduced a new algorithm based on the ideas of Kamiya and Talman (1990). They improved already existing algorithms to solve a stationary point problem on a polytope like the one of Yamamoto (1987). Applying the latter

algorithm to our problem would imply that in cases 2 and 3 of our *SLSPP*-algorithm described in Section 2 we would be obliged to calculate a sequence of vertices of S_A^n. This means solving a sequence of linear programming problems similar to the case when we had to show that the starting point was an end point of the path followed by the algorithm. The fact that Kamiya and Talman (1990) presented an alternative avoiding the calculation of vertices of S_A^n but using affinely independent points instead made application of their ideas to our equilibrium problem very worthwile.

Appendix

Lemma 1 *Let $w^0, w^1, \ldots, w^m$ be affinely independent points in $\mathbf{R}^{n+1}$. Given $\delta_i \in \mathbf{R}$, $i = 0, 1, \ldots, m$, define the points $\hat{w}^0, \hat{w}^1, \ldots, \hat{w}^m$ as*

$$\hat{w}^i = w^i + \delta^i(q^1 - q^0),$$

for $i = 0, 1, \ldots, m$, where q^1 and q^0 are $(n+1)$-vectors such that for certain numbers λ_h^0 and λ_h^1, $h \in \mathcal{I}_m \cup \{0\}$, $q^1 \neq q^0$,

$$q^0 = \sum_{h=0}^{m} \lambda_h^0 w^h \text{ and } \sum_{h=0}^{m} \lambda_h^0 = 1,$$

and

$$q^1 = \sum_{h=0}^{m} \lambda_h^1 w^h \text{ and } \sum_{h=0}^{m} \lambda_h^1 = 1.$$

Let $g \in \{0, 1, \ldots, m\}$ be such that $\lambda_g^0 \neq \lambda_g^1$. Then the points $\hat{w}^i$, $i = 0, 1, \ldots, g-1$, $g+1, \ldots, m$, are affinely independent.

Proof: Let β_i, $i = 0, 1, \ldots, g-1, g+1, \ldots, m$, be such that

$$\sum_{i=0,\ i\neq g}^{m} \beta_i \hat{w}^i = 0, \qquad \sum_{i=0,\ i\neq g}^{m} \beta_i = 0.$$

Then we have to prove that $\beta_i = 0$ for all $i \neq g$. Substituting $\hat{w}^i$ into this expression gives

$$\sum_{i=0,\ i\neq g}^{m} \beta^i w^i + \left(\sum_{i=0,\ i\neq g}^{m} \beta_i \delta^i \right) \left(\sum_{h=1}^{m} (\lambda_h^1 - \lambda_h^0)(w^h - w^0) \right) = 0.$$

Defining δ as

$$\delta = \sum_{i=0,\ i\neq g}^{m} \beta_i \delta^i,$$

this expression reduces to

$$\sum_{i=0,\ i\neq g}^{m} \beta_i w^i + \delta \sum_{h=1}^{m} (\lambda_h^1 - \lambda_h^0)(w^h - w^0) = 0.$$

Rewriting this result in a suitable way gives

$$\sum_{i=1,\ i\neq g}^{m} (\beta_i + \delta(\lambda_i^1 - \lambda_i^0))w^i + (\beta_0 - \delta \sum_{h=1}^{m} (\lambda_h^1 - \lambda_h^0))w^0 + \delta(\lambda_g^1 - \lambda_g^0)w^g = 0.$$

Let

$$\begin{aligned}
\gamma_i &= \beta_i + \delta(\lambda_i^1 - \lambda_i^0) \text{ for all } i \neq 0, g, \\
\gamma_0 &= \beta_0 - \delta \sum_{h=1}^{m} (\lambda_h^1 - \lambda_h^0), \\
\gamma_g &= \delta(\lambda_g^1 - \lambda_g^0).
\end{aligned}$$

Then

$$\sum_{i=0}^{m} \gamma_i = \sum_{i=0,\ i\neq g}^{m} \beta_i = 0$$

and

$$\sum_{i=0}^{m} \gamma_i w^i = 0.$$

By affine independence of $w^0, w^1, \ldots, w^m$ it follows that $\gamma_i = 0$, $i = 0, 1, \ldots, m$. Therefore

$$\begin{aligned}
&\beta_i + \delta(\lambda_i^1 - \lambda_i^0) = 0 \text{ for all } i \neq 0, g, \\
&\beta_0 - \delta \textstyle\sum_{h=1}^{m} (\lambda_h^1 - \lambda_h^0) = 0, \\
&\delta(\lambda_g^1 - \lambda_g^0) = 0.
\end{aligned}$$

We took $g \in \{0, 1, \ldots, m\}$ such that $\lambda_g^1 - \lambda_g^0 \neq 0$. Therefore $\delta = 0$. But then $\beta_i = 0$ for all $i \neq g$. □

Chapter 10

A simplicial variable dimension restart algorithm on S_A^n

The previous two chapters of this monograph were concerned with the computation of an equilibrium in an economy with linear production technologies in a very efficient way. The algorithm introduced in Chapter 9 seems to be the most efficient one. The main drawback of the algorithms introduced in these two chapters is the impossibility to attach any economic interpretation to the adjustment process to be followed by the algorithm and that the sequence of points generated by these algorithms can not be shown to converge to an equilibrium. This chapter aims at introducing an algorithm that incorporates a plausible economic interpretation of the underlying adjustment process. Moreover it generates a sequence of approximating equilibria, in the sense that any convergent subsequence converges to an equilibrium. It however lacks the efficiency of the *SLSPP*-algorithm introduced in Chapter 9 as it incorporates in an economy with $n+1$ commodities and l production technologies a pivot system of $n+l+2$ equations instead of $n+1$ equations and it is not able to start on the boundary of the feasible set.

The algorithm introduced in this chapter follows approximately the path originating from the adjustment process introduced in van den Elzen, van der Laan, and Talman (1990). This adjustment process is described in Section 1. Section 2 introduces the simplicial variable dimension restart algorithm to approximate the path resulting from this adjustment process. This piecewise linear path either leads to an

approximating equilibrium in an economy with linear production technologies or it diverges towards infinity. It can be shown that the path generated by the algorithm is bounded and therefore always ends up with an approximating equilibrium.

10.1 The adjustment process

In van den Elzen, van der Laan, and Talman (1990) an adjustment process has been introduced to find an equilibrium in an economy with linear production technologies as defined in Definition 8.1.1. This adjustment process can be initiated in an arbitrarily chosen point p^0 in the (relative) interior of the set S_A^n defined in Chapter 9 and generates under some regularity conditions a piecewise smooth path of prices and activity levels connecting the initial price vector with an equilibrium. Along this path the prices and activity levels are adjusted in a very simple and natural way.

Let (p^*, y^*) constitute an equilibrium in an economy with $n+1$ commodities and l linear production technologies. Then the combination of the conditions imposed on (p^*, y^*) in Definition 8.1.1 and the properties that $p^{*\mathsf{T}} A y^* = 0$ and $p^{*\mathsf{T}}(z(p^*) - Ay^*) = 0$ proved in Property 8.1.2 and Property 8.1.3, respectively, implies that

$$\begin{array}{lll} \text{if} & z_j(p^*) - (Ay^*)_j < 0 & \text{then} \quad \frac{p_j^*}{p_j^0} = \min_h \frac{p_h^*}{p_h^0} = 0, \\ \text{if} & z_j(p^*) - (Ay^*)_j = 0 & \text{then} \quad \frac{p_j^*}{p_j^0} \geq \min_h \frac{p_h^*}{p_h^0}, \\ \text{if} & p^{*\mathsf{T}} a^i < 0 & \text{then} \quad y_i^* = 0, \\ \text{if} & p^{*\mathsf{T}} a^i = 0 & \text{then} \quad y_i^* \geq 0, \end{array} \tag{10.1}$$

for $j \in \mathcal{I}_{n+1}$ and $i \in \mathcal{I}_l$. This interpretation of the conditions to hold at an equilibrium in an economy with linear production technologies makes it very natural at p^0 to decrease p_j proportionally from p_j^0 if $z_j(p^0) < 0$ and to increase p_i proportionally from p_i^0 if $z_i(p^0) > 0$. Without loss of generality we assume that $z_j(p^0) \neq 0$ for every $j \in \mathcal{I}_{n+1}$. Because of Walras' law there always exists an index $j \in \mathcal{I}_{n+1}$ such that $z_j(p^0) < 0$ and an index $i \in \mathcal{I}_{n+1}$ such that $z_i(p^0) > 0$. This gives for the adjustment process $2^{n+1} - 2$ possible rays to leave p^0. Given these rays to leave p^0 the adjustment process is such that it generates a path of points (p, y) in $S_A^n \times \mathbf{R}_+^l$ where each point

(p, y) satisfies the following conditions

$$\begin{array}{llll} \text{if} & z_j(p) - (Ay)_j < 0 & \text{then} & \min_h \frac{p_h}{p_h^0} = \frac{p_j}{p_j^0}, \\ \text{if} & z_j(p) - (Ay)_j = 0 & \text{then} & \min_h \frac{p_h}{p_h^0} \leq \frac{p_j}{p_j^0} \leq \max_h \frac{p_h}{p_h^0}, \\ \text{if} & z_j(p) - (Ay)_j > 0 & \text{then} & \frac{p_j}{p_j^0} = \max_h \frac{p_h}{p_h^0}, \\ \text{if} & p^\top a^i < 0 & \text{then} & y_i = 0, \\ \text{if} & p^\top a^i = 0 & \text{then} & y_i \geq 0, \end{array} \tag{10.2}$$

for $i \in \mathcal{I}_l$ and $j \in \mathcal{I}_{n+1}$. If $z(p^*) - Ay^* \leq 0$ for some price vector p^* and some vector of activity levels y^* on the path then in (p^*, y^*) the conditions in (10.2) reduce to the conditions in (10.1) and therefore (p^*, y^*) is an end point of the path yielding an equilibrium in the economy.

As proved in van den Elzen (1991) the pairs of price vector $p \in S_A^n$ and activity level $y \in \mathbf{R}_+^l$ satisfying (10.2) form piecewise smooth paths and loops. Each path is bounded because of Assumption 8.1.2, saying that there is no production without input in the economy. Hence each path has therefore two end points. Exactly one end point is the point $(p, y) = (p^0, 0)$ whereas all other end points constitute equilibria. Therefore there is one path connecting $(p^0, 0)$ with an equilibrium while every other path connects two equilibria. The adjustment process now generates the path connecting $(p^0, 0)$ with an equilibrium. Clearly $(p^0, 0)$ satisfies (10.2) for both the maximum and the minimum equal to one. Initially the prices of the commodities for which there is an excess demand are proportionally increased from its minimum, those for which there is an excess supply are proportionally decreased from its maximum, and there is no production since all firms make per unit activity losses. As soon as the demand for a commodity becomes equal to its supply then either an equilibrium has been found or its price is relatively increased if there was excess supply before and relatively decreased if there was excess demand before while demand and supply of this commodity are kept equal to each other. In this way the price ratio of a commodity in excess demand (excess supply) is kept maximal (minimal) with respect to the other price ratios.

When along the path the price ratio of a commodity in equilibrium becomes relatively maximal (minimal) then the commodity becomes in excess demand (excess supply) while its price ratio is kept relatively maximal (minimal).

Finally, when an activity starts to make zero profits along the path then its activity level is increased from zero while keeping its profits equal to zero. Also when the activity level of an activity becomes zero along the path then its profit is made negative while keeping its activity level equal to zero.

10.2 The algorithm

The algorithm to follow the piecewise smooth path generated by the adjustment process as described in the previous section is a simplicial algorithm on $S^n \times \mathbf{R}^l_+$. First the set S^n is subdivided into simplices and the excess demand function z is linearized on each simplex of the subdivision. With respect to the piecewise linear approximation on this subdivision the path becomes piecewise linear and connects the starting point with an approximating equilibrium. Each linear piece of the path is followed by making a linear programming pivot step in a system of $n+l+2$ linear equations. Since the number of simplices is finite, the set of solutions to each system is bounded, and since no cycling can occur, the algorithm terminates within a finite number of pivot steps. When the accuracy of approximation is not sufficient the algorithm can be restarted at the approximating equilibrium with a finer simplicial subdivision in order to increase the accuracy.

Let $s \in \mathbf{R}^{n+1}$ be a sign vector. Then we can define a subset $A(s)$ of S^n as follows.

Definition 10.2.1 *For each sign vector $s \in \mathbf{R}^{n+1}$ the subset $A(s)$ of S^n is given by*

$$A(s) = \emptyset \text{ if } \mathcal{I}^{+1}(s) = \emptyset \text{ or } \mathcal{I}^{-1}(s) = \emptyset,$$

otherwise

$$A(s) = \{p \in S^n \mid \tfrac{p_j}{p_j^0} = \max_h \tfrac{p_h}{p_h^0} \text{ for } j \in \mathcal{I}^{+1}(s),$$
$$\tfrac{p_j}{p_j^0} = \min_h \tfrac{p_h}{p_h^0} \text{ for } j \in \mathcal{I}^{-1}(s)\}.$$

Figure 10.2.1 gives a subdivision of S^n into subsets $A(s)$ when $n = 2$. A nonempty set $A(s)$ is a t-dimensional subset of S^n, where $t = \mid \mathcal{I}^0(s) \mid +1$. Clearly, a point $(p, y) \in S^n_A \times \mathbf{R}^l_+$ not being an equilibrium satisfies (10.2) if and only if there exists

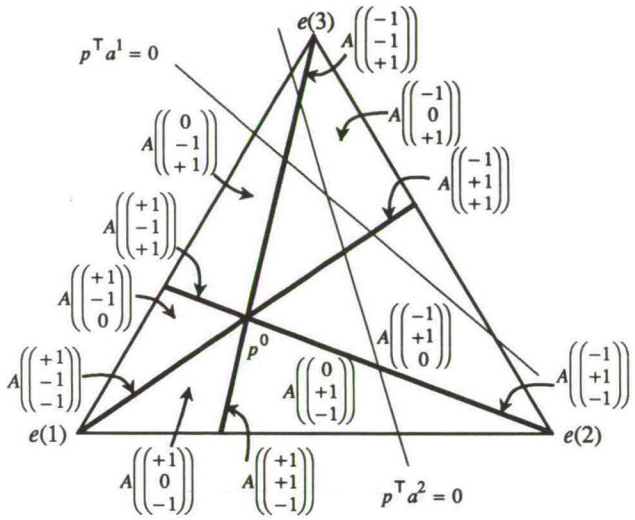


FIGURE 10.2.1: Subdivision of S^2 into subsets $A(s)$ for sign vectors $s \in \mathbf{R}^3$.

a unique sign vector s and some subset $\mathcal{U}$ of $\mathcal{I}_l$ such that

$$\begin{aligned} &1) \quad p \in A(s) \text{ and } \operatorname{sgn}(z(p) - Ay) = s \\ &2) \quad p^T a^i = 0 \text{ and } y_i \geq 0 \text{ for all } i \in \mathcal{U} \\ &\qquad p^T a^i \leq 0 \text{ and } y_i = 0 \text{ for all } i \notin \mathcal{U}. \end{aligned} \tag{10.3}$$

Let $\mathcal{G}$ be a simplicial subdivision of S^n such that it subdivides each nonempty set $A(s)$ into t-dimensional simplices. A specific simplicial subdivision that can easily be implemented is the V-triangulation, see Doup, van der Laan, and Talman (1987). The excess demand function z is then linearized on $\mathcal{G}$ by taking its piecewise linear approximation with respect to $\mathcal{G}$, denoted by Z. When we replace in (10.2) the function z by its piecewise linear approximation Z then the set of points $(p, y) \in S^n_A \times \mathbf{R}^l_+$ satisfying (10.2) with respect to Z consists of piecewise linear paths and loops, one path connecting $(p^0, 0)$ with an end point $(\overline{p}, \overline{y})$. This end point $(\overline{p}, \overline{y})$ can be considered as an approximating equilibrium. Also all other end points of the other existing paths are approximating equilibria.

The piecewise linear path connecting $(p^0, 0)$ with an approximating equilibrium $(\overline{p}, \overline{y})$ consists of points (p, y) satisfying the conditions in (10.3) with z replaced by Z.

Let t be the dimension of $A(s)$. If $p \in A(s)$ then there are a t-simplex $\sigma(u^1, \ldots, u^{t+1})$ of $\mathcal{G}$ in $A(s)$ and numbers $\lambda_1, \ldots, \lambda_{t+1} \geq 0$ such that $p = \sum_{i=1}^{t+1} \lambda_i u^i$ and $\sum_{i=1}^{t+1} \lambda_i = 1$, i.e. p is contained in $\sigma(u^1, \ldots, u^{t+1})$. On the other hand, if $\operatorname{sgn}(Z(p) - Ay) = s$, then there exist $\mu_h \geq 0$, $h \notin \mathcal{I}^0(s)$, such that

$$Z(p) - Ay = \sum_{h \notin \mathcal{I}^0(s)} \mu_h s_h \bar{e}(h),$$

where $\bar{e}(h)$ denotes the $(n+1)$-dimensional h-th unit vector. Furthermore the second condition in (10.3) implies the existence of $\nu_i \geq 0$, $i \notin \mathcal{U}$, such that

$$A^\top p + \sum_{i \notin \mathcal{U}} \nu_i \underline{e}(i) = 0,$$

where $\underline{e}(i)$ denotes the l-dimensional i-th unit vector. If one combines these results with the conditions in (10.3) then (p, y) satisfies (10.2) if and only if for some sign vector $s \in \mathbf{R}^{n+1}$ and subset $\mathcal{U}$ of $\mathcal{I}_l$ the system of linear equations

$$\sum_{j=1}^{t+1} \lambda_j \begin{pmatrix} z(u^j) \\ A^\top u^j \\ 1 \end{pmatrix} - \sum_{i \in \mathcal{U}} y_i \begin{pmatrix} a^i \\ 0 \\ 0 \end{pmatrix} - \sum_{j \notin \mathcal{I}^0(s)} \mu_j s_j \begin{pmatrix} \bar{e}(j) \\ 0 \\ 0 \end{pmatrix} + \sum_{i \notin \mathcal{U}} \nu_i \begin{pmatrix} 0 \\ \underline{e}(i) \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{10.4}$$

corresponding to the t-simplex $\sigma(u^1, \ldots, u^{t+1})$ in $A(s)$ has a solution $\lambda_j \geq 0$ $(j \in \mathcal{I}_{t+1})$, $y_i \geq 0$ $(i \in \mathcal{U})$, $\mu_j \geq 0$, $(j \notin \mathcal{I}^0(s))$, $\nu_i \geq 0$ $(i \notin \mathcal{U})$ such that $p = \sum_{j=1}^{t+1} \lambda_j u^j$. This system contains $n + l + 2$ equations with $n + l + 3$ unknowns leaving us with one degree of freedom. Assuming nondegeneracy, the set of solutions to system (10.4) if nonempty and bounded represents a line segment. This line segment corresponds to a linear piece of points in $A(s, \mathcal{U})$ where

$$A(s, \mathcal{U}) := \{(p, y) \in (S_A^n \cap A(s)) \times \mathbf{R}_+^l \mid p^\top a^i = 0 \text{ for } i \in \mathcal{U} \text{ and } y_i = 0 \text{ for } i \notin \mathcal{U}\}.$$

This line segment can be followed by the algorithm by making a linear programming pivot step in (10.4) with the column vector corresponding to the variable being zero in one of its end points. In the other end point of this line segment, say $(\overline{\lambda}, \overline{y}, \overline{\mu}, \overline{\nu})$, either $\overline{\lambda}_g = 0$ for some $g \in \mathcal{I}_{t+1}$, or $\overline{y}_r = 0$ for some $r \in \mathcal{U}$, or $\overline{\mu}_k = 0$ for some $k \notin \mathcal{I}^0(s)$, or $\overline{\nu}_f = 0$ for some $f \notin \mathcal{U}$. Let $\overline{p} = \sum_{i=1}^{t+1} \overline{\lambda}_i u^i$.

Case 1: $\overline{\lambda}_g$ is zero for some $g \in \mathcal{I}_{t+1}$. Then $\overline{p}$ lies in the facet τ of σ opposite the vertex y^g. This facet τ is either a facet of exactly one other t-simplex, say $\overline{\sigma}$, in $A(s)$ or τ lies in the boundary of $A(s)$.

Suppose τ is a facet of exactly one other t-simplex $\overline{\sigma}$ in $A(s)$. Then, in order to continue the path in $A(s,\mathcal{U})$, a pivot step is made in (10.4) with the column $(z(\overline{u})^\top, \overline{u}^\top A, 1)^\top$ corresponding to the unique vertex $\overline{u}$ of $\overline{\sigma}$ not contained in τ.

Suppose τ lies in the boundary of $A(s)$. Then either $\overline{p}$ lies in the face $S^n(\mathcal{I}^{-1}(s)) := \{p \in S^n \mid p_i = 0,\ i \in \mathcal{I}^{-1}(s)\}$ of S^n or $\overline{p} \in A(\overline{s})$ where $\overline{s}$ is a sign vector such that $\overline{s}_j \neq 0$ for some $j \in \mathcal{I}^0(s)$ and $\overline{s}_i = s_i$ for all $i \neq j$. If $\overline{p} \in S^n(\mathcal{I}^{-1}(s))$ then $\overline{p}_h = 0$ for all $h \in \mathcal{I}^{-1}(s)$ and $\overline{p}_h > 0$ for all $h \in \mathcal{I}^{+1}(s)$. Moreover $Z_h(\overline{p}) - (A\overline{y})_h \leq 0$ for all h for which $\overline{p}_h = 0$ and $Z_h(\overline{p}) - (A\overline{y})_h \geq 0$ for all h for which $\overline{p}_h > 0$. Hence $\overline{p}$ is an approximating solution to the equilibrium problem and the algorithm terminates.

If $\overline{p} \in A(\overline{s})$ with $\overline{s}$ defined as above then $(\overline{p}, \overline{y})$ is an end point of a linear piece of the path in $A(\overline{s},\mathcal{U})$. This linear piece can be followed by pivoting the column $-\overline{s}_j(\overline{e}(j)^\top, 0, 0)^\top$ into (10.4) thereby raising μ_j from zero.

Case 2: $\overline{y}_r$ is zero for some $r \in \mathcal{U}$. Then $(\overline{p}, \overline{y})$ is an end point of a linear piece of the path in $A(s,\mathcal{U} \setminus \{r\})$. This linear piece can be followed by pivoting the column $(0, \underline{e}(r)^\top, 0)^\top$ into (10.4) thereby raising ν_r from zero.

Case 3: $\overline{\mu}_k$ is zero for some $k \notin \mathcal{I}^0(s)$. Let $\overline{s}$ be a sign vector such that $\overline{s}_k = 0$ and $\overline{s}_h = s_h$ for $h \neq k$. If $\overline{s} \leq 0$ or $\overline{s} \geq 0$ then $Z(\overline{p}) - A\overline{y} \leq 0$ or $Z(\overline{p}) - A\overline{y} \geq 0$. Hence $(\overline{p}, \overline{y})$ is an approximating equilibrium and the algorithm terminates. Otherwise the end point $(\overline{p}, \overline{y})$ is an end point of a linear piece of the path in $A(\overline{s},\mathcal{U})$. In $A(\overline{s})$ there is exactly one $(t+1)$-simplex $\overline{\sigma}$ having σ as a facet. Then the algorithm follows the linear piece in $A(\overline{s},\mathcal{U})$ by pivoting the column $(z(\overline{u})^\top, \overline{u}^\top A, 1)^\top$ into (10.4), where $\overline{u}$ is the vertex of $\overline{\sigma}$ not contained in σ.

Case 4: $\overline{\nu}_f$ is zero for some $f \notin \mathcal{U}$ and so $\overline{p}^\top a^f = 0$. Then $(\overline{p}, \overline{y})$ is an end point of a linear piece on the path in $A(s,\mathcal{U} \cup \{f\})$. This linear piece is generated by the algorithm by pivoting the column $(-a^{f\top}, 0, 0)^\top$ into the system of equations (10.4) thereby raising y_f from zero.

The algorithm starts in $(p^0, 0)$ where $p^0 \in \text{int}(S^n_A)$. To show that $(p^0, 0)$ is an end point to a unique line segment of solutions to (10.4) denote $\text{sgn}(z(p^0))$ by s^0. Because we assume nondegeneracy s^0 does not contain any components equal to zero. Hence $z(p^0) = \sum_{j=1}^{n+1} \mu_j^0 s^0 \overline{e}(j)$ for unique $\mu_j^0 > 0$, $j \in \mathcal{I}_{n+1}$. Furthermore, $p^0 \in \text{int}(S^n_A)$ implies that $p^{0\top} A + \sum_{i=1}^{l} \nu_i^0 \underline{e}(i) = 0$ for unique $\nu_i^0 > 0$, $i \in \mathcal{I}_l$. Hence, at $(p^0, 0)$ the

system

$$\lambda_1 \begin{pmatrix} z(p^0) \\ A^\top p^0 \\ 1 \end{pmatrix} - \sum_{j=1}^{n+1} \mu_j s_j^0 \begin{pmatrix} \overline{e}(j) \\ 0 \\ 0 \end{pmatrix} + \sum_{i=1}^{l} \nu_i \begin{pmatrix} 0 \\ \underline{e}(i) \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \tag{10.5}$$

has a unique solution $\lambda_1 = 1$, $\mu_j = s_j^0 z_j(p^0) > 0$, $j \in \mathcal{I}_{n+1}$, and $\nu_i = -a^{i\top} p^0 > 0$, $i \in \mathcal{I}_l$. The point p^0 is a vertex of a unique 1-simplex $\sigma(u^1, u^2)$ in $A(s^0)$ with $u^1 = p^0$. The algorithm therefore starts with pivoting the column $(z(u^2)^\top, (u^2)^\top A, 1)^\top$ into (10.5) thereby generating a linear piece of the path in $A(s^0, \emptyset)$.

Since all steps are unique and returning to $(p^0, 0)$ is impossible the algorithm either terminates within a finite number of steps with an approximating equilibrium $(\overline{p}, \overline{y})$ or it diverges towards infinity. The cases in which an approximating equilibrium is found are summarized in the next lemma.

Lemma 10.2.1 *Let $(\overline{p}, \overline{y})$ be an end point of a linear piece of the path corresponding to a line segment of solutions to* (10.4) *for some sign vector $s \in \mathbf{R}^{n+1}$. Then $(\overline{p}, \overline{y})$ is an approximating equilibrium in the economy with linear production technologies if at the solution $(\overline{\lambda}, \overline{y}, \overline{\mu}, \overline{\nu})$ one of the following cases holds:*

i) $\overline{\lambda}_g = 0$ for some $g \in \mathcal{I}_{t+1}$ and $\overline{p} \in S^n(\mathcal{I}^{-1}(s))$;

ii) $\overline{\mu}_k = 0$ for some $k \notin \mathcal{I}^0(s)$, and either $\mathcal{I}^{+1}(s) = \{k\}$ or $\mathcal{I}^{-1}(s) = \{k\}$.

When an approximating equilibrium $(\overline{p}, \overline{y})$ is found one can measure the accuracy of approximation by taking the smallest $\epsilon > 0$ for which for $j \in \mathcal{I}_{n+1}$

$$\begin{aligned} z_j(\overline{p}) - (A\overline{y})_j \leq \epsilon \quad \text{if} \quad \overline{p}_j = 0, \\ -\epsilon \leq z_j(\overline{p}) - (A\overline{y})_j \leq \epsilon \quad \text{if} \quad \overline{p}_j > 0. \end{aligned} \tag{10.6}$$

If $z(\overline{p}) - A\overline{y}$ is not accurate enough, i.e., ϵ is too large, the algorithm is repeated with a finer simplicial subdivision of S^n being started at or close to $(\overline{p}, 0)$ such that the new p^0 lies in the interior of S_A^n. This in the hope to find a more accurate approximation within a few number of steps.

The algorithm would not end up in an approximating equilibrium if it diverges towards infinity. In van den Elzen (1990) it has been proved that because of Assumption 8.1.2 the path obtained from the adjustment process is bounded and hence

always ends up with an equilibrium. A similar proof can be given for the piecewise linear path generated by the algorithm if one replaces the continuous excess demand function z in the proof in van den Elzen (1990) by its piecewise linear approximation.

Bibliography

[1] Brouwer, L.E.J. (1912), "Über Abbildung von Mannigfaltigkeiten", *Mathematische Annalen 71*, 97-115.

[2] Cottle, R.W. (1966), "Nonlinear programs with positively bounded Jacobians", *SIAM Journal of Applied Mathematics 14*, 147-158.

[3] Dai, Y., and Talman, A.J.J. (1990), "Linear stationary point problems on unbounded polyhedra", CentER Discussion Paper 9067, Center for Economic Research, Tilburg University, Tilburg, The Netherlands, to appear in Mathematics of Operations Research.

[4] Dai, Y., Laan, G. van der, Talman, A.J.J. and Yamamoto, Y. (1991), "A simplicial algorithm for the nonlinear stationary point problem on an unbounded polyhedron", *SIAM Journal of Applied Mathematics 91*, 151-165.

[5] Dang, C. (1991a), "The D_1-triangulation of $\mathbf{R}^n$ for simplicial algorithms for computing solutions of nonlinear equations", *Mathematics of Operations Research 16*, 148-161.

[6] Dang, C. (1991b), *The D_1-Triangulation in Simplicial Algorithms*, Ph.D.-thesis, Tilburg University, Tilburg, The Netherlands.

[7] Dantzig, G.B. (1951), "Maximization of a linear function of variables subject to linear inequalities", in: T.C. Koopmans, ed., *Activity Analysis of Production and Allocation,* Wiley, New York, NY, USA, pp. 339-347.

[8] Debreu, G. (1959), *Theory of Value,* Yale University Press, New Haven, CT, USA.

[9] Doup, T.M. (1988), *Simplicial Algorithms on the Simplotope*, Springer, Berlin, Germany.

[10] Doup, T.M., Laan, G. van der, and Talman, A.J.J. (1987), "The $(2^{n+1}-2)$-ray algorithm: A new simplicial algorithm to compute economic equilibria", *Mathematical Programming 39*, 241-252.

[11] Doup, T.M., and Talman, A.J.J. (1987), "A new variable dimension simplicial algorithm to find equilibria on the product space of unit simplices", *Mathematical Programming 37*, 319-355.

[12] Eaves, B.C. (1972), "Homotopies for the computation of fixed points", *Mathematical Programming 3*, 1-22.

[13] Eaves, B.C. (1987), "Thoughts on computing market equilibrium with $SLCP$", in: Talman A.J.J., and Laan, G. van der, eds., *The Computation and Modelling of Economic Equilibria*, North-Holland, Amsterdam, The Netherlands, pp. 1-18.

[14] Elzen, A.H. van den (1991), *Adjustment Processes for Exchange Economies and Noncooperative Games*, Ph.D.-thesis, Tilburg University, Tilburg, The Netherlands.

[15] Elzen, A.H. van den, Laan, G. van der, and Talman, A.J.J. (1990), "An adjustment process for an exchange economy with linear production technologies", CentER Discussion Paper 9015, Center for Economic Research, Tilburg University Tilburg, The Netherlands.

[16] Elzen, A.H. van den, Kremers, J.A.W.M., Laan, G. van der, and Talman, A.J.J. (1991), "A simplicial algorithm to compute an equilibrium in an economy with linear production technologies", mimeo, Department of Econometrics, Tilburg University, Tilburg, The Netherlands.

[17] Evers, J.J.M. (1978), "More with the Lemke complementarity algorithm", *Mathematical Programming 15*, 214-219.

[18] Fisher, M.L., and Gould, F.J. (1974), "A simplicial algorithm for the nonlinear complementarity problem", *Ibid. 6*, 281-300.

[19] Freudenthal, H. (1942), "Simplizialzerlegungen von beschränkter Flachheit", *Ann. of Math. 43*, 580-582.

[20] Garcia, C.B. (1973), *The Complementarity Problem and its Applications*, Ph.D.-thesis, Rensselaer Polytechnic Institute, Troy, NY, USA.

[21] Hansen, T. (1968), *On the Approximation of a Competetive Equilibrium*, Ph.D.-thesis, Yale University, New Haven, CT, USA.

[22] Jones, P.C. (1986), "Even more with the Lemke complementarity algorithm", *Mathematical Programming 25*, 239-242.

[23] Kakutani, S. (1941), "A generalization of Brouwer's fixed point theorem", *Duke Mathematical Journal 8*, 457-459.

[24] Karush, W. (1939), *Minima of Functions of Several Variables with Inequalities as Side Constraints*, M.Sc. Dissertation, Department of Mathematics, University of Chicago, Chicago, IL, USA.

[25] Kamiya, K., and Talman, A.J.J. (1990), "Linear stationary point problems", CentER Discussion Paper 9022, Center for Economic Research, Tilburg University, Tilburg, The Netherlands.

[26] Kojima, M. (1974), "Computational methods for solving the nonlinear complementarity problem", *Keio Engrg. Rep. 27 No.1*, 1-41.

[27] Kremers, J.A.W.M., and Talman, A.J.J. (1990a), "Solving the nonlinear complementarity problem with lower and upper bounds", *Methods of Operations Research 60*, 407-417.

[28] Kremers, J.A.W.M., and Talman, A.J.J. (1990b), "Solving the nonlinear complementarity problem", *Methods of Operations Research 62*, 91-103.

[29] Kremers, J.A.W.M., and Talman, A.J.J. (1990c), "A new algorithm for the linear complementarity problem allowing for an arbitrary starting point", FEW 464, Department of Econometrics, Tilburg University, Tilburg, The Netherlands, to appear in Mathematical Programming.

[30] Kremers, J.A.W.M., and Talman, A.J.J. (1991), "An *SLSPP*-algorithm to compute an equilibrium in an economy with linear production technologies", FEW 496, Department of Econometrics, Tilburg University, Tilburg, The Netherlands.

[31] Kuhn, H.W. (1960), "Some combinatorial lemma's in topology", *I.B.M. Journal of Research Development 4*, 518-524.

[32] Kuhn, H.W., and Tucker, A.W. (1951), "Nonlinear Programming", in: Metropolis, N.C., ed., *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, 1950, University of California Press, Berkeley, CA, USA, pp.481-492 [reprinted in 1968 in: Newman, P., ed., *Readings in Mathematical Economics, Vol.I, Value Theory*, The Johns Hopkins Press, Baltimore, MA, USA, pp. 3-14.]

[33] Laan, G. van der (1980), *Simplicial Fixed Point Algorithms,* Mathematical Centre Tracts 129, Mathematisch Centrum, Amsterdam, The Netherlands.

[34] Laan, G. van der, and Talman, A.J.J. (1979), "A restart algorithm for computing fixed points without an extra dimension", *Mathematical Programming 17*, 74-84.

[35] Laan, G. van der, and Talman, A.J.J. (1981), " A class of simplicial restart fixed point algorithms without an extra dimension", *Mathematical Programming 20*, 33-48.

[36] Laan, G. van der, and Talman, A.J.J. (1985), "An algorithm for the linear complementarity problem with upper and lower bounds", *Journal of Optimization Theory and Applications 62 No.1*, 151-163.

[37] Laan, G. van der, and Talman, A.J.J. (1987), "Simplicial approximation of solutions to the nonlinear complementarity problem with lower and upper bounds", *Mathematical Programming 38*, 1-15.

[38] Lemke, C.E. (1965), "Bimatrix equilibrium points and mathematical programming", *Management Science 11*, 681-689.

[39] Lemke, C.E., and Howson Jr., J.T. (1964), "Equilibrium points of bimatrix games", *SIAM Journal of Applied Mathematics 12*, 413-423.

[40] Mathiesen, L. (1985a), "Computational experience in solving equilibrium models by a sequence of linear complementarity problems", *Operations Research 33*, 1225-1250.

[41] Mathiesen, L. (1985b), "Computation of economic equilibria by a sequence of linear complementarity problems", *Mathematical Programming Study 23*, 144-162.

[42] Mathiesen, L. (1987), "An algorithm based on a sequence of linear complementarity problems applied to a Walrasian equilibrium model: an example", *Mathematical Programming 37*, 1-18.

[43] Mathiesen, L., and Rutherford, T. (1983), "Testing the Robustness of an Iterative *LCP* Algorithm for Solving Walrasian Equilibrium Models", Discussion Paper 0883, Norwegian School of Economics and Business Administration, Bergen, Norway.

[44] Merrill, O.H. (1971), *Applications and Extensions of an Algorithm that Computes Fixed Points of Certain Upper Semi-continuous Point-to-Set Mappings*, Ph.D.-thesis, University of Michigan, Ann Arbor, MI, USA.

[45] Murty, K.G. (1988), *Linear Complementarity, Linear and Nonlinear Programming*, Sigma Series in Applied Mathematics, Heldermann, Berlin, Germany.

[46] Scarf, H. (1960), "Some examples of global instability of the competetive equilibrium", *International Economic Review 1 No.3*, 157-172.

[47] Scarf, H. (1967), "The approximation of fixed points of a continuous mapping", *SIAM Journal of Applied Mathematics 15*, 1328-1343.

[48] Scarf, H.E. (1973), with the collaboration of T. Hansen, *Computation of Economic Equilibria,* Yale University Press, New Haven, CT, USA.

[49] Schrijver, A. (1986), *Theory of Linear and Integer Programming*, Wiley, New York, NY, USA.

[50] Sperner, E. (1928), "Neuer Beweis für die Invarianz der Dimensionszahl und des Gebietes", *Abh. Math. Sem. Univ. Hamburg 6*, 265-272.

[51] Talman, A.J.J. (1980), *Variable Dimension Fixed Point Algorihms and Triangulations,* Mathematical Centre Tracts 128, Mathematisch Centrum, Amsterdam, The Netherlands.

[52] Talman, A.J.J., and Van der Heyden, L. (1983), "Algorithms for the linear complementarity problem which allow an arbitrary starting point", in: B.C. Eaves et all., eds., *Homotopy Methods and Global Convergence*, Plenum Press, New York, NY, USA, pp. 267-285.

[53] Talman, A.J.J., and Yamamoto, Y. (1989), "A simplicial algorithm for stationary point problems on polytopes", *Mathematics of Operations Research 14*, 383-399.

[54] Todd, M.J. (1976), *The Computation of Fixed Points and Applications*, Springer, Berlin, Germany.

[55] Todd, M.J. (1978), "Improving the convergence of fixed point algorithms", *Mathematical Programming Study 7*, 151-179.

[56] Walras, L. (1874), *Eléménts d'Économie Politique Pure,* Corbaz, Lausanne, Switzerland.

[57] Wright, A.H. (1981), "The octahedral algorithm, a new simplicial fixed point algorithm", *Mathematical Programming 21*, 47-69.

[58] Yamamoto, Y. (1987), "A path following algorithm for stationary point problems", *Journal of the OR Society of Japan 30*, 181-198.

Samenvatting

Dit proefschrift heeft twee doelstellingen. Ten eerste worden algoritmes geïntroduceerd ter berekening van een evenwicht in een ruileconomie met lineaire productietechnologieën. Ten tweede worden algoritmes gepresenteerd voor het vinden van een oplossing voor complementariteitsproblemen, bekend uit de mathematische programmering. Deze twee toepassingen zijn sterk aan elkaar gerelateerd daar de meeste economische evenwichtsproblemen de vorm hebben van een bepaald complementariteitsprobleem.

Een ruileconomie met lineaire productietechnologieën bestaat uit een eindig aantal goederen, consumenten en productie-activiteiten. De consumenten hebben een bepaalde voorraad in elk van deze goederen. Bovendien oefenen zij een bepaalde vraag uit naar deze goederen gegeven de prijzen van deze goederen. De producenten bepalen de hoeveelheid input of output van elk goed in de economie door een bepaald niveau van de productie-activiteiten vast te stellen. Een evenwicht in een economie met lineaire productietechnologieën wordt gegeven door een stelsel prijzen van goederen en activiteitenniveau's zodanig dat aan de vraag naar elk goed in de economie voldaan kan worden door het aanbod van dit goed door de producenten en de consumenten, en zodanig dat geen van de productie-activiteiten een positieve winst maakt. Eén van de belangrijkste bijdragen aan de berekening van een evenwicht in een ruileconomie met lineaire productietechnologieën is het door Mathiesen in 1985 geïntroduceerde algoritme. Mathiesen stelde voor om het stelsel evenwichtsvoorwaarden in een ruileconomie met lineaire productietechnologieën te benaderen door een rij van lineaire complementariteitsproblemen (*SLCP*).

Allereerst normaliseert Mathiesen de prijzen in de economie op de eenheidssimplex. Dit betekent dat alle prijzen tot één optellen. Daarna koos hij een willekeurige prijsvector uit deze eenheidssimplex als startpunt voor het *SLCP*-algoritme. In deze willekeurig gekozen prijsvector berekent Mathiesen de zogenaamde eerste-orde Taylorbenadering van de vraagoverschotfunctie in de economie. Daardoor wordt het niet-lineaire stelsel van evenwichtsvoorwaarden omgeschreven tot een lineair stelsel.

De homogeniteit van de graad nul van de vraagfunctie van de consumenten leidt ertoe dat de matrix met eerste-orde afgeleiden van deze vraagfunctie, de Jacobiaan, singulier is. Het probleem van deze singulariteit kan op twee manieren opgelost worden. Men kan een normaliseringsbeperking aan de prijzen opleggen. Dit betekent in dat men bijvoorbeeld de prijzen op de eenheidssimplex neemt. Deze voorwaarde wordt dan aan het gelineariseerde stelsel toegevoegd. Een andere manier om aan de singulariteitsproblemen tegemoet te komen is om een bepaald goed als numéraire aan te duiden. In dat geval stelt men de prijs van het betreffende geod vast op een willekeurig gekozen waarde en berekent de evenwichtsprijzen van de overige goederen in de betreffende economie gegeven de prijs van het numéraire goed. Door Mathiesen werd gekozen voor de laatste oplossing.

Het gelineariseerde stelsel heeft de vorm van een lineair complementariteitsprobleem. Door de toepassing van het *LCP*-algoritme van Lemke uit 1965 was Mathiesen

niet altijd in staat om een oplossing voor het lineaire complementariteitsprobleem te berekenen, zelfs al had het betreffende lineaire complementariteitsprobleem een oplossing. Indien, gegeven de keuze van de numéraire, er geen oplossing gevonden kon worden, stelde Mathiesen voor om gewoon een andere prijs als numéraire te nemen. Deze procedure wordt net zo lang herhaald totdat er een lineair complementariteitsprobleem gevonden is waarvoor het Lemke-algoritme wel in staat is om een oplossing te berekenen. De gevonden prijsvector wordt op de eenheidssimplex geprojecteerd en deze projectie dient dan als startpunt voor de volgende iteratie in het *SLCP*. Op deze manier wordt mogelijk een rij prijsvectoren op de eenheidsimplex verkregen die mogelijkerwijs convergeert naar een evenwichtsoplossing.

Eaves stelde in 1987 voor om het probleem van de singuliere Jacobiaan op te lossen door een normaliseringsbeperking op de prijzen toe te voegen aan het stelsel gelineariseerde evenwichtsvoorwaarden. Eaves toont daartoe aan dat het evenwichtsprobleem in een economie met lineaire productietechnologieën equivalent is met het stationaire punt probleem van de vraagoverschotfunctie op een deelverzameling van de eenheidssimplex. Deze deelverzameling volgt uit de doorsnede van de condities die door de normaliseringsvoorwaarde aan de prijzen wordt opgelegd en de condities resulterende uit de "geen winst"-voorwaarde op de evenwichtsprijzen. Dit stationaire punt probleem werd door Eaves zodanig herschreven dat het betreffende probleem opgelost kan worden door een reeks van lineaire complementariteitsproblemen in de trant van Mathiesen. Volgens Eaves leent zijn *SLCP* zich beter voor een theoretische analyse dan Mathiesen's *SLCP*. Daarenboven toont Eaves aan dat zijn *SLCP* het lineaire complementariteitsprobleem, dat in elke iteratie verkregen wordt, oplost. Dit in tegenstelling tot Mathiesen.

In 1990 introduceren Kamiya en Talman een bijzonder efficiënt algoritme voor de berekening van een stationair punt van een affiene functie op een polytoop. Zoals eerder vermeld is het evenwichtsprobleem in een economie met lineaire productietechnologieën equivalent met het stationair punt probleem op een deelverzameling van de eenheidssimplex, zijnde een polytoop. Daarom stellen wij voor om het evenwichtsprobleem door een rij lineaire stationaire punt problemen te benaderen en elk van deze problemen op te lossen door een algoritme gebaseerd op de ideeën van Kamiya en Talman. Op deze wijze lijkt een efficiënter alternatief te zijn verkregen voor de *SLCP*-algoritmes van Eaves en Mathiesen.

De algoritmes om het evenwichtsprobleem in een economie met lineaire productietechnologieën op te lossen door het probleem te benaderen met een sequentieel algoritme lijken erg snel te zijn indien ze in staat zijn om een convergerende rij benaderende oplossingen te berekenen. Het grootste nadeel van deze algoritmes is echter het ontbreken van globale convergentie. Om aan dit nadeel tegemoet te komen introduceren we een simpliciaal variabele dimensie algoritme om het evenwichtsprobleem op te lossen. Dit algoritme benadert het pad dat geïnduceerd wordt door het aanpassingsproces dat in 1990 ontwikkeld werd door Van den Elzen, Van der Laan en Talman. Het algoritme lineariseert stuksgewijs de vraagoverschotfunctie op

een simpliciale opdeling van de eenheidssimplex. Het kan bewezen worden dat dit algoritme een benaderend evenwicht vindt. Indien men niet tevreden is met de nauwkeurigheid van de verkregen benadering dan kan de procedure herstart worden met een fijnere simpliciale opdeling van de eenheidssimplex. In tegenstelling tot bovengenoemde sequentiële algoritmes kan dit algoritme nogal traag zijn indien men een goede benadering van het evenwicht wil bereiken. Daar staat echter tegenover dat het aanpassingsproces wat door het algoritme gevolgd wordt van een economische interpretatie voorzien kan worden.

Zoals reeds eerder vermeld neemt het evenwichtsprobleem in een ruileconomie met lineaire productietechnologieën de vorm aan van een bepaald complementariteitsprobleem. Deze complementariteitsproblemen komen frequent voor in de mathematische programmering alsook in andere onderzoeksgebieden waar optimalisering een belangrijke rol speelt en vormen het onderwerp van een ander deel van dit proefschrift. Het meest eenvoudige en tevens het meest bekende complementariteitsprobleem is het reeds eerder vermelde lineaire complementariteitsprobleem. Het betreft hier het vinden van niet-negatieve vectoren die onderling affien en complementair zijn. Het lineaire complementariteitsprobleem omvat bekende problemen als lineaire programmering, kwadratische programmering en evenwichtsproblemen in bimatrix spelen als een speciaal geval. Verder komt men dit probleem vaak tegen bij het oplossen van Karush-Kuhn-Tucker condities bij optimalisering. Het bekendste algoritme dat in staat is om een oplossing voor het lineaire complementariteitsprobleem te vinden is het complementaire pivoting algoritme geïntroduceerd door Lemke in 1965 en Lemke en Howson omstreeks 1964. Een nadeel van dit algoritme is echter dat het algoritme alleen in de oorsprong kan starten. Dit geeft problemen bij een gevoeligheidsanalyse en het leidt tot inefficiëntie bij het herstarten in een *SLCP*. In 1983 introduceerden Talman en Van der Heyden een klasse van algoritmes dat het Lemke-algoritme zodanig veralgemeniseerd dat het mogelijkheid is om in een willekeurig gekozen punt te starten. In dit proefschrift wordt een algoritme voor het lineaire complementariteitsprobleem met de mogelijkheid van een willekeurig gekozen startpunt gepresenteerd dat efficiënter is en een meer natuurlijke interpretatie heeft dan de algoritmes geïntroduceerd van Talman en Van der Heyden.

Cottle introduceerde in 1966 het niet-lineaire complementariteitsprobleem waarbij de affiene relatie tussen de variabelen vervangen is door een willekeurige continue relatie. De evenwichtsproblemen in een zuivere ruileconomie alsmede in een ruileconomie met lineaire productietechnologieën hebben de vorm van een niet-lineair complementariteitsprobleem. De algoritmes zoals die door Mathiesen in 1985 geïntrodu- ceerd werden kunnen daarom ook gebruikt worden om een oplossing voor het niet-lineaire complementariteitsprobleem te vinden. Begin zeventiger jaren introduceren Kojima, Fisher en Gould, Garcia, en Merrill bijvoorbeeld algoritmes om een oplossing voor het niet-lineaire complementariteitsprobleem te berekenen. Al deze methoden zijn ontstaan als een variant op het algoritme dat door Scarf in 1967 gebruikt werd om op een constructieve wijze de bekende dekpuntstelling van Brouwer uit 1912 te bewijzen.

In dit proefschrift introduceren wij een simpliciaal variabele dimensie algoritme ter berekening van een oplossing van het niet-lineaire complementariteitsprobleem.

Het niet-lineaire complementariteitsprobleem is wederom een speciaal geval van het niet-lineaire complementariteitsprobleem met onder- en bovengrenzen. Dit probleem staat ook wel bekend als het gegeneralizeerde niet-lineaire complementariteitsprobleem. Van der Laan en Talman introduceerden in 1985 een algoritme voor de berekening van een oplossing in het lineaire geval terwijl in zij in 1987 het niet-lineaire geval opgelosten. In hun algoritme voor het niet-lineaire geval gebruiken Van der Laan en Talman de K'-triangulering welke door Todd in 1978 geïntroduceerd werd. In dit proefschrift wordt de door Doup en Talman in 1987 geïntroduceerde V-triangulering gebruikt.

Dit proefschrift bestaat uit drie delen. Deel A is de introductie tot het proefschrift. In hoofdstuk 2 wordt de wiskundige ondergrond behandeld welke noodzakelijk is voor het begrip van de overige hoofdstukken. Dit hoofdstuk bevat enkele basisstellingen en resultaten uit de lineaire algebra met betrekking tot polytopen en kegels en het toont de relatie tussen dekpunten, stationaire punten en complementariteit. In hoofdstuk 3 wordt de uit de lineaire programmeringstheorie bekende pivotprocedure besproken alsmede het concept van een simpliciale opdeling van een convexe verzameling. Verder wordt in dit hoofdstuk een raamwerk voor de in dit proefschrift geïntroduceerde algoritmes gegeven.

In Deel B worden algoritmes voor bekende complementariteitsproblemen geïntroduceerd. Hoofdstuk 4 beschrijft het Lemke algoritme alsmede een enigszins aangepaste versie van één van de algoritmes geïntroduceerd door Talman en Van der Heyden in 1983. Daarbij maken we gebruik van het in hoofdstuk 3 geïntroduceerde raamwerk. Verder worden in dit hoofdstuk enkele bekende problemen uit de mathematische programmering bekeken die een speciaal geval van een lineair complementariteitsprobleem zijn. Hoofdstuk 5 introduceert een nieuw algoritme ter berekening van een oplossing voor het lineaire complementariteitsprobleem. Hoofdstuk 6 van dit proefschrift beschouwt het niet-lineaire complementariteitsprobleem met onder- en bovengrenzen en introduceert een simpliciaal variabele dimensie herstart algoritme ter berekening van een oplossing voor dit probleem. Hoofdstuk 7 breidt dit algoritme uit naar het niet-lineaire complementariteitsprobleem.

Deel C van dit proefschrift concentreert zich op algoritmes ter berekening van een evenwicht in een economie met lineaire productietechnologieën. Het concept van een economie met lineaire productietechnologieën alsmede de *SLCP*-algoritmes van Mathiesen en van Eaves worden beschreven in hoofdstuk 8. Hoofdstuk 9 beschrijft de rij van lineaire stationaire punt problemen ter benadering van het evenwichtsprobleem in een economie met lineaire productietechnologieën als een alternatief voor de algoritmes uit hoofdstuk 8. Hoofdstuk 10 introduceert het simpliciale variabele dimensie herstart algoritme om het aanpassingsproces zoals beschreven door Van den Elzen, Van der Laan en Talman te benaderen.